

WO 2006/006528

1

10/579128

PCT/JP2005/012667

明 細 書  
12 MAY 2006

音響信号検出システム、音響信号検出サーバ、映像信号探索装置、映像信号探索方法、映像信号探索プログラムと記録媒体、信号探索装置、信号探索方法、及び、信号探索プログラムと記録媒体

#### 技術分野

[0001] 本発明は、蓄積された蓄積音響信号の中から、この蓄積音響信号よりも短い乃至同じ長さの目的音響信号と類似した信号の位置を探し出す信号検出に関するものであり、例えば実環境中の音響信号検出に用いられる音響信号検出システム及び音響信号検出サーバに関する。

すなわち、本発明は、実環境中に流れている音楽やCM(Commercial Message:コマーシャル)を携帯端末で受信し、その受信された音響信号を用いて膨大な音楽CMデータベースの中から同一の音楽やCMを検索するものである。

[0002] また、本発明は、実環境下で取得した特徴ひずみを含む映像信号(目的映像信号)と類似の映像信号をデータベース中に蓄積された映像信号(蓄積映像信号)から探索する映像信号探索装置、映像信号探索方法、映像信号探索プログラム及び記録媒体に関する。

[0003] また、本発明は、実環境下で取得した特徴ひずみを含む信号(目的信号)と類似の信号をデータベース中に蓄積された信号(蓄積信号)から探索する信号探索装置、信号探索方法、信号探索プログラム及び記録媒体に関する。

[0004] 本願は、2004年7月9日に出願された特願2004-203198号に対し優先権を主張し、その内容をここに援用する。

本願は、2004年7月9日に出願された特願2004-203199号に対し優先権を主張し、その内容をここに援用する。

本願は、2005年3月17日に出願された特願2005-077726号に対し優先権を主張し、その内容をここに援用する。

本願は、2005年3月22日に出願された特願2005-082107号に対し優先権を主張し、その内容をここに援用する。

WO 2006/006528

2

PCT/JP2005/012667

本願は、2005年3月24日に出願された特願2005-086280号に対し優先権を主張し、その内容をここに援用する。

## 背景技術

[0005] 近年、音や映像など、マルチメディア情報の流通が盛んに行われるようになり、このマルチメディア情報を得るための検索や探索が必要となってきた。

例えば、実環境中に流れている音楽や映像またはCMを携帯端末などにより受信し、その受信された信号(目的音響信号または目的映像信号)を用いて、膨大なデータベースの中から同一の音楽や映像、CMを検索することが考えられる。

このため、具体的に探索したい音や映像の信号を目的音響信号あるいは目的映像信号として指定して、これに類似する蓄積音響信号あるいは蓄積映像信号がデータベースのいずれに存在するかの探索を行う必要がある。

時系列信号の高速一致法として、例えば、特許第3065314号が開示されている。ここで行われる探索は、時系列探索であり、高速かつ高精度に行われる必要がある。

[0006] ところが、ユーザが実環境において収録して入手する音楽や映像には、様々な特性ひずみ、例えば、スピーカやスクリーンなどの信号発生源の機器の特性、入力する携帯端末の特性による乗法性ひずみや、実環境中の要素(音であれば雑音、反響、吸収など、映像であれば大気透明度、明度角度ゆれ、反射など)による加法性雑音が含まれていることが考えられる。これに対し、特許第3065314号公報による方法は、検索対象となる信号について特徴ひずみが少ないことが前提となっており、雑音やひずみが含まれる場合においてはその探索精度は著しく低下する。

[0007] この欠点を解決するため、入力される信号に対し変動を付加する変動付加過程を設けることで、特徴ひずみに対して頑健な信号検出を行う方法が提案されている(例えば、特許第3408800号公報参照)。ところが、上述した変動付加過程を設ける際、複数の雑音やひずみを考慮する必要がある場合、これらの複数の目的特徴を用意しなければならず、情報量が大幅に増大してしまうという欠点がある。

[0008] また、入力される音や映像の目的信号の強度ピークを検出し、このピークにおける周波数を特徴とすることにより、雑音やひずみに対して頑健な信号検出を行う方法が

WO 2006/006528

3

PCT/JP2005/012667

提案されている。

- [0009] しかしながら、この方法には、入力される信号のピークを用いるため、この実際の信号のピーク周辺にある大きな雑音の影響を受けることにより、実際の信号のピーク検出に失敗し探索精度が低下するという欠点がある。
- [0010] このため、局所的な特徴の統計量を用いて入力される音や映像の目的信号を正規化することにより、ひずみに対して頑健な信号検出を行う方法が提案されている(例えば、特開2003-022084号公報や、特に音響信号に関してはWO 02/11123 A2号公報参照)。すなわちこの信号検出の方法は、特徴ひずみによる音や映像の目的信号の変動を吸収するため、周波数特徴を抽出した後、時間一周波数空間上の局所領域毎に目的信号を正規化して、特徴ひずみに頑健な空間へのデータ変換を行い、この空間において音や映像の蓄積音響信号との比較を行うというものである。
- [0011] しかしながら、上述した特徴ひずみに頑健な空間へのデータ変換を行う方法においては、雑音や途切れ、非定常なひずみに対し、正規化だけではその特徴ひずみを吸収することができず、探索の信頼性を低下させるという欠点がある。
- [0012] 本発明はこのような事情に鑑みてなされたものであり、検索対象である音や映像の信号(目的信号: 目的音響信号および目的映像信号)を加法的雑音や途切れに対して頑健なデータに変換して探索精度を向上させ、かつ探索時間を大幅に短縮させることを目的とする。
- [0013] また、乗法的ひずみを含む信号に対し、これを頑健なデータに変換することにより探索精度を向上し、高速高精度の処理を実行することが可能なシステムを提供することを目的とする。
- [0014] また、線形量子化よりも量子化後の値の信頼性を高めるような量子化手段を適用することで、全体の探索精度を向上させることができる信号探索装置、信号探索方法、信号探索プログラム及び記録媒体を提供することを目的とする。

## 発明の開示

- [0015] 上記目的を達成するために、本発明の音響信号検出システムは、蓄積音響信号か

WO 2006/006528

4

PCT/JP2005/012667

ら、該蓄積音響信号よりも短い乃至同じ長さの目的音響信号に類似した部分を探し出す信号検出システムにおいて、前記蓄積音響信号の時系列データから、特徴ベクトルからなる蓄積特徴を計算する蓄積特徴計算部と、前記目的音響信号の時系列データから、特徴ベクトルからなる目的特徴を計算する目的特徴計算部と、前記蓄積特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、該蓄積特徴から選択し、該選択された要素のベクトルからなる蓄積エリア選択特徴を導く蓄積特徴エリア選択部と、前記目的特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、該目的特徴から選択し、該選択された要素のベクトルからなる目的エリア選択特徴を導く目的特徴エリア選択部と、前記蓄積エリア選択特徴中に照合区間を設定し、前記目的エリア選択特徴及び前記蓄積エリア選択特徴中の該照合区間のそれぞれとの類似度を計算する特徴照合部とを有し、前記特徴照合部が前記蓄積エリア選択特徴において照合区間を順次移動させて繰り返し、前記類似度から目的エリア選択特徴と類似する蓄積エリア選択特徴の領域を探索することを特徴とする。

[0016] 特に、「高速信号検出法、装置およびその記録媒体」(特許第3065314号公報)及び「信号検出方法、装置及びそのプログラム、記録媒体」(特許第3408800号公報)に比較して、新たに蓄積特徴エリア選択部及び目的特徴エリア選択部を設けて、特徴ベクトルの特に特徴のある要素を選択して、この選択した要素のみを比較しているため、雑音・とぎれに対して特許第3065314号公報及び2に比較して頑健な音響信号検出を可能とした。

[0017] また、「信号検出方法及び装置、プログラムならびに記録媒体」(特開2003-022084号公報)の目的信号及び蓄積信号を正規化し、単純に比較判定する方法に対して、新たに蓄積特徴エリア選択部及び目的特徴エリア選択部を設けて、特徴ベクトルの特に特徴のある要素を選択して、この選択した要素のみを比較しているため、雑音・とぎれに対して上記手法に比較して頑健な音響信号検出を可能とした。

さらに、「System and methods for recognizing sound and music signal in high noise and distortion」(WO 02/11123 A2号公報)に比較して、新たに蓄積特徴エリア選択部及び目的特徴エリア選択部を設けて、特徴ベクトルの特に



WO 2006/006528

5

PCT/JP2005/012667

特徴のある要素を選択して、この選択した要素のみを比較しているため、雑音・とぎれに対して上記手法に比較して頑健な音響信号検出を可能とした。

[0018] 本発明によれば、雑音や途切れの影響を除去するため、蓄積音響信号および目的音響信号から統計的に特徴的な要素のみを選択し照合を行う。この選択された要素からなる多次元ベクトルを用いることで、蓄積音響信号と目的音響信号のより特徴的なパターンを比較してそれらの類似度を算出することができ、無駄な部分の比較処理を行わないため、目的音響信号に重畳している雑音や音響の途切れの影響を大幅に低減させ、頑健な音響信号検出を行うことができる。

[0019] また、実際に蓄積音響信号及び目的音響信号から得られる蓄積特徴及び目的特徴から、統計的に評価してより特徴的な要素からなる蓄積エリア選択特徴及び目的エリア選択特徴を抽出することで、照合するデータ量を大幅に低減することが可能となり、類似度の計算処理を高速化するとともに、蓄積する蓄積音響信号の1件あたりのデータ量を大幅に削減することができ、従来と同様の記憶部容量に、より多くの蓄積音響信号情報を蓄積することが可能となり、この点からも類似した蓄積音響信号の検出精度を向上させることができる。

[0020] また、上記目的を達成するために、本発明の音響信号検出システムは、蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの目的音響信号に類似した部分を探し出す信号検出システムにおいて、前記蓄積音響信号の時系列データから、特徴ベクトルからなる蓄積特徴を計算する蓄積特徴計算部と、前記目的音響信号の時系列データから、特徴ベクトルからなる目的特徴を計算する目的特徴計算部と、前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所定の統計量を計算し、蓄積特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化部と、前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化部と、前記蓄積正規化特徴から、量子化して得た値を要素とするベクトルからなる蓄積量子化特徴を導く蓄積特徴量子化部と、前記目的正規化特

WO 2006/006528

6

PCT/JP2005/012667

徴から、量子化して得た値を要素とするベクトルからなる目的量子化特徴を導く目的特徴量子化部と、前記蓄積量子化特徴中に照合区間を設定し、前記目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算する特徴照合部とを有し、前記特徴照合部が前記蓄積量子化特徴において照合区間を順次移動させて繰り返し、前記類似度から目的量子化特徴と類似する蓄積量子化特徴の領域を探索することを特徴とする。

[0021] 特に、「高速信号検出法、装置およびその記録媒体」(特許第3065314号公報)及び「信号検出方法、装置及びそのプログラム、記録媒体」(特許第3408800号公報)に比較して、新たに目的特徴正規化部、蓄積特徴正規化部、目的特徴量子化部及び蓄積特徴量子化部を設けて特徴ベクトルの各要素を、正規化した後、所定の閾値により量子化することで、ひずみに対して特許第3065314号公報及び2に比較して頑健な音響信号検出を可能とした。

[0022] また、「信号検出方法及び装置、プログラムならびに記録媒体」(特開2003-022084号公報)の目的信号及び蓄積信号を正規化し、単純に比較判定する方法に加えて、新たに目的特徴量子化部及び蓄積特徴量子化部を設け、特徴ベクトルの各要素を量子化することにより、特徴ひずみに対して上記手法に比較して頑健な音響信号検出を可能とした。

さらに、「System and methods for recognizing sound and music signal in high noise and distortion」(WO 02/11123 A2号公報)に比較して、新たに目的特徴正規化過程及び蓄積特徴正規化過程を設けて特徴を正規化することにより、特徴ひずみに対して上記手法に比較して頑健な音響信号検出を可能とした。

[0023] 本発明によれば、特徴ひずみを吸収するため、所定の閾値により音響信号をスカラー量子化する量子化部(過程)を設けることにより、比較するデータ量を大幅に低減することが可能となり、類似度の計算処理を高速化するとともに、蓄積する蓄積音響信号の1件あたりのデータ量を大幅に削減することができ、従来と同様の記憶部容量により多くの蓄積音響信号情報を蓄積することができ、この点からも類似した蓄積音響信号の検出精度を向上させることができる。

WO 2006/006528

7

PCT/JP2005/012667

- [0024] また、本発明によれば、上記量子化部(過程)を設けることにより、各要素毎のデータの詳細な比較判定でなく、照合区間における特徴ベクトルを全体的なパターンとして検索するので、探索精度を向上させて、検出洩れを防ぐことが可能となり、公知の方法と比較して、さまざまな特徴ひずみに頑健な信号検出処理が行え、より汎用的な特徴ひずみに頑健な音響信号検出を行うことができる。
- [0025] さらに、上記問題を解決するために、本発明は、蓄積映像信号の中から目的映像信号に類似した信号を探索する映像信号探索装置であって、目的映像信号から目的特徴を計算する目的特徴計算手段と、目的特徴から目的統計量を計算する目的統計量計算手段と、目的統計量に対して所定の閾値により閾値処理し、目的統計量を選択し、選択された目的統計量を要素とするベクトル又は行列からなる目的エリア選択特徴を算出する目的特徴エリア選択手段と、蓄積特徴から所定の蓄積統計量を計算する蓄積統計量計算手段と、蓄積統計量に対して所定の閾値により閾値処理し、蓄積統計量を選択し、選択された蓄積統計量を要素とするベクトル又は行列からなる蓄積エリア選択特徴を算出する蓄積特徴エリア選択手段と、蓄積エリア選択特徴に対して照合区間を設定し、照合区間中の蓄積エリア選択特徴と目的エリア選択特徴の少なくとも一部との類似度を計算する特徴照合手段とを有し、特徴照合手段において、照合区間を順次移動させて繰り返し類似度を計算することを特徴とする。
- [0026] 本発明は、上記に記載の発明において、前記目的特徴エリア選択部及び前記蓄積特徴エリア選択部は、各々、統計量として目的特徴及び蓄積特徴の、第1の所定時間区間における平均値を算出し、前記要素から該平均値を減算した値の絶対値が所定の閾値を越えるものを選択するようにしたことを特徴とする。
- [0027] 本発明は、上記に記載の発明において、前記目的特徴量算出手段及び前記蓄積統計量計算手段は、各々、目的特徴及び蓄積特徴の要素毎に第2の所定時間区間における平均値と標準偏差とを算出し、該平均値と標準偏差とを用いて、目的特徴及び蓄積特徴を正規化することにより、目的統計量及び蓄積統計量を算出することを特徴とする。
- [0028] また、本発明は、蓄積映像信号の中から目的映像信号に類似した信号を探索する映像信号探索方法であって、目的映像信号から目的特徴を計算する目的特徴計算

WO 2006/006528

8

PCT/JP2005/012667

工程と、目的特徴から目的統計量を計算する目的統計量計算工程と、目的統計量に対して所定の閾値により閾値処理し、目的統計量を選択し、選択された目的統計量を要素とするベクトル又は行列からなる目的エリア選択特徴を算出する目的特徴エリア選択工程と、蓄積特徴から所定の蓄積統計量を計算する蓄積統計量計算工程と、蓄積統計量に対して所定の閾値により閾値処理し、蓄積統計量を選択し、選択された蓄積統計量を要素とするベクトル又は行列からなる蓄積エリア選択特徴を算出する蓄積特徴エリア選択工程と、蓄積エリア選択特徴に対して照合区間を設定し、照合区間中の蓄積エリア選択特徴と目的エリア選択特徴の少なくとも一部との類似度を計算する特徴照合工程とを有し、特徴照合工程において、照合区間を順次移動させて繰り返し類似度を計算することを特徴とする映像信号探索方法である。

[0029] また、本発明は、上記に記載の発明のいずれか1つに記載の映像信号探索装置としてコンピュータを機能させるための映像信号探索プログラムである。

[0030] また、本発明は、上記に記載の発明のコンピュータプログラムを記録したコンピュータ読み取り可能な記録媒体である。

[0031] 本発明によれば、反射や途切れの影響を除去するため、蓄積映像信号及び目的映像信号から統計的に特徴的な要素のみを選択し照合を行う。この選択された要素からなる多次元ベクトルを用いることで、蓄積映像信号と目的映像信号のより特徴的なパターンを比較してそれらの類似度を算出することができ、無駄な部分の比較処理を行わないため、目的映像信号に重畳している反射や映像の途切れの影響を大幅に低減させ、雑音に頑健な映像信号検出を行うことができる。

[0032] また、蓄積映像信号及び目的映像信号から得られる蓄積特徴及び目的特徴から、統計的に評価してより特徴的な要素からなる蓄積エリア選択特徴及び目的エリア選択特徴を抽出することで、照合するデータ量を大幅に低減することが可能となり、類似度の計算処理を高速化するとともに、蓄積する蓄積映像信号の1件あたりのデータ量を大幅に削減することができ、従来と同様の記憶容量部に、より多くの蓄積映像信号情報を蓄積することができる。

[0033] つまり、蓄積特徴エリア選択部と目的特徴選択部を設けることで、目的映像中に含まれる、反射により物体が写りこんだ箇所やキャプチャに失敗して途切れた箇所を回

WO 2006/006528

9

PCT/JP2005/012667

避して、目的映像信号と蓄積映像信号との特徴量同士の照合を行うことができる。これにより、エリア選択せずに照合する場合に比べて、相対的な類似度が上昇し、探索精度を向上することができる。

[0034] さらに、上記問題を解決するために、本発明は、蓄積映像信号の中から目的映像信号に類似した信号を探索する装置であって、目的映像信号から目的特徴を計算する目的特徴計算手段と、目的特徴から目的統計量を計算する目的統計量計算手段と、目的統計量と目的特徴とを用いて、目的正規化特徴の要素を算出する目的特徴正規化手段と、目的正規化特徴の要素を所定の閾値を用いて、量子化し、目的量子化特徴の要素を求め、目的ベクトルを生成する目的量子化手段と、蓄積特徴から蓄積統計量を計算する蓄積統計量計算手段と、蓄積統計量と蓄積特徴とを用いて、蓄積正規化特徴の要素を算出する蓄積特徴正規化手段と、蓄積正規化特徴の要素を所定の閾値を用いて、量子化し、蓄積量子化特徴の要素を求め、蓄積ベクトルを生成する蓄積量子化手段と、蓄積ベクトルに対して照合区間を設定し、照合区間中の蓄積ベクトルの要素と目的ベクトルの少なくとも一部の要素との類似度を計算する特徴照合手段とを有し、特徴照合手段において、照合区間を順次移動させて繰り返し類似度を計算することを特徴とする映像信号探索装置である。

[0035] また、本発明は、蓄積映像信号の中から目的映像信号に類似した信号を探索する方法であって、目的映像信号から目的特徴を計算する目的特徴計算工程と、目的特徴から目的統計量を計算する目的統計量計算工程と、目的統計量と目的特徴とを用いて、目的正規化特徴の要素を算出する目的特徴正規化工程と、目的正規化特徴の要素を所定の閾値を用いて、量子化し、目的量子化特徴の要素を求め、目的ベクトルを生成する目的量子化工程と、蓄積特徴から蓄積統計量を計算する蓄積統計量計算工程と、蓄積統計量と蓄積特徴とを用いて、蓄積正規化特徴の要素を算出する蓄積特徴正規化工程と、蓄積正規化特徴の要素を所定の閾値を用いて、量子化し、蓄積量子化特徴の要素を求め、蓄積ベクトルを生成する蓄積量子化工程と、蓄積ベクトルに対して照合区間を設定し、照合区間中の蓄積ベクトルの要素と目的ベクトルの少なくとも一部の要素との類似度を計算する特徴照合工程とを有し、特徴照合工程において、照合区間を順次移動させて繰り返し類似度を計算することを特

WO 2006/006528

10

PCT/JP2005/012667

徴とする映像信号探索方法である。

- [0036] また、本発明は、上記に記載の映像信号探索装置としてコンピュータを機能させるための映像信号探索プログラムである。
- [0037] また、本発明は、上記に記載の映像信号プログラムを記録したコンピュータ読み取り可能な記録媒体である。
- [0038] 本発明によれば、特徴ひずみを吸収するため、所定の閾値により映像信号をスカラー量子化する量子化部を設けることにより、比較するデータ量を大幅に低減することが可能となり、類似度の演算処理を高速化するとともに、蓄積する蓄積映像信号の1件あたりのデータ量を大幅に削減することができ、従来と同様の記憶部容量により多くの蓄積映像信号を蓄積することができ、この点からも類似した蓄積映像信号の検出精度を向上させることができる。
- [0039] また、本発明によれば、量子化部を設けることにより、各要素のデータの詳細な比較判定でなく、照合区間における特徴ベクトルを全体的なパターンとして検索するので、探索精度を向上させて、検出洩れを防ぐことが可能となり、従来の方法と比較して、さまざまな特徴ひずみに頑健な信号検出処理が行え、より汎用的な特徴ひずみに頑健な映像信号検出を行うことができる。
- [0040] さらに、上記問題を解決するために、本発明は、蓄積信号の中から目的信号に類似した信号を探索する信号探索装置であって、目的信号から目的特徴を計算する目的特徴計算手段と、目的特徴から目的統計量を計算する目的統計量計算手段と、目的統計量と目的特徴とを用いて、目的正規化特徴の要素を算出する目的特徴正規化手段と、目的正規化特徴の要素を入力し、選択された目的非線形量子化特徴の要素を求め、目的ベクトルを生成する目的エリア選択非線形量子化手段と、蓄積特徴から蓄積統計量を計算する蓄積統計量計算手段と、蓄積統計量と蓄積特徴とを用いて、蓄積正規化特徴の要素を算出する蓄積特徴正規化手段と、蓄積正規化特徴の要素を入力とし、選択された蓄積非線形量子化特徴の要素を求め、蓄積ベクトルを生成する蓄積エリア選択非線形量子化手段と、蓄積ベクトルに対して照合区間を設定し、照合区間中の蓄積ベクトルの要素と目的ベクトルの少なくとも一部の要素との類似度を計算する特徴照合手段とを有し、特徴照合手段において、照合区間

WO 2006/006528

11

PCT/JP2005/012667

を順次移動させて繰り返し類似度を計算することを特徴とする。

- [0041] 本発明は、上記に記載の発明において、目的エリア選択非線形量子化手段及び蓄積エリア選択非線形量子化手段は、多次元ベクトルをボロノイ分割し、特徴ベクトルが属するボロノイ境界面との距離を非線形量子化することを特徴とする。
- [0042] 本発明は、上記に記載の発明において、目的エリア選択非線形量子化手段は、統計量が所定の閾値を超える要素を選択し、選択された要素の多次元ベクトルからなる目的エリア選択特徴を導く目的特徴エリア選択手段と、特徴ベクトルを非線形量子化する目的特徴非線形量子化手段とからなることを特徴とする。
- [0043] 本発明は、上記に記載の発明において、蓄積エリア選択量子化手段は、統計量が所定の閾値を超える要素を選択し、選択された要素の多次元ベクトルからなる蓄積エリア選択特徴を導く蓄積特徴エリア選択手段と、特徴ベクトルを非線形量子化する蓄積特徴非線形量子化手段とからなることを特徴とする。
- [0044] 本発明は、蓄積信号の中から目的信号に類似した信号を探索する信号探索方法であって、目的信号から目的特徴を計算する目的特徴計算工程と、目的特徴から目的統計量を計算する目的統計量計算工程と、目的統計量と目的特徴とを用いて、目的正規化特徴の要素を算出する目的特徴正規化工程と、目的正規化特徴の要素を入力し、選択された目的非線形量子化特徴の要素を求め、目的ベクトルを生成する目的エリア選択非線形量子化工程と、蓄積特徴から蓄積統計量を計算する蓄積統計量計算工程と、蓄積統計量と蓄積特徴とを用いて、蓄積正規化特徴の要素を算出する蓄積特徴正規化工程と、蓄積正規化特徴の要素を入力し、選択された蓄積非線形量子化特徴の要素を求め、蓄積ベクトルを生成する蓄積エリア選択非線形量子化工程と、蓄積ベクトルに対して照合区間を設定し、照合区間中の蓄積ベクトルの要素と目的ベクトルの少なくとも一部の要素との類似度を計算する特徴照合工程とを有し、特徴照合工程において、照合区間を順次移動させて繰り返し類似度を計算することを特徴とする。
- [0045] また、本発明は、上記に記載の信号探索装置としてコンピュータを機能させるための信号探索プログラムである。
- [0046] また、本発明は、上記に記載の信号探索プログラムを記録したコンピュータ読み取

WO 2006/006528

12

PCT/JP2005/012667

り可能な記録媒体である。

[0047] また、本発明によれば、雑音やひずみの影響を除去するため、蓄積信号及び目的信号から統計的に特徴的な要素のみを抽出し照合を行う。この抽出された要素からなる多次元ベクトルを用いることで、蓄積信号と目的信号のより特徴的なパターンを認識した類似度計算が行え、更に、統計的处理により算出される統計量に非線形量子化を適用することで量子化後の値の信頼性を向上させ、目的信号に重畳している雑音やひずみの影響を低減させながら高精度の信号検出が可能となる。

[0048] また、統計的に評価してより特徴的な要素からなる蓄積エリア選択特徴及び目的エリア選択特徴を抽出し、更に、蓄積エリア選択特徴及び目的エリア選択特徴を非線形量子化している。これにより、照合するデータ量を大幅に低減することが可能となり、類似度の計算処理を高速化するとともに、蓄積する蓄積信号の1件あたりのデータ量を大幅に削減することができる。つまり、従来と同様の記憶容量部に、より多くの蓄積信号情報を蓄積することが可能となり、この点からも類似した蓄積信号の検出精度を向上させることができる。

#### 図面の簡単な説明

[0049] [図1]図1は、本発明の一実施例による音響信号検出システムの構成例を示すブロック図である。

[図2]図2は、図1の音響信号検出システムの動作例を示すフローチャートである。

[図3]図3は、本発明の一実施例による音響信号検出システムの構成例を示すブロック図である。

[図4]図4は、図3の音響信号検出システムの動作例を示すフローチャートである。

[図5]図5は、本発明の一実施例による映像信号探索システムの構成例を示すブロック図である。

[図6]図6は、本発明の一実施例による映像信号探索システムにおける閾値の設定の説明に用いるグラフである。

[図7]図7は、本発明の一実施例による映像信号探索システムにおける閾値の設定の説明に用いるグラフである。



WO 2006/006528

13

PCT/JP2005/012667

[図8]図8は、図5の映像信号探索システムの動作例を示すフローチャートである。

[図9]図9は、本発明の一実施例による映像信号探索システムにおけるサブ画面の説明図である。

[図10]図10は、本発明の一実施例による映像信号探索システムに局所時間とフレームの関係の説明図である。

[図11]図11は、本発明の一実施形態による映像信号探索システムの構成例を示すブロック図である。

[図12]図12は、図11の映像信号探索システムの動作例を示すフローチャートである。

[図13]図13は、本発明の一実施形態による映像信号探索システムにおける時間窓とフレームの関係の説明図である。

[図14]図14は、本発明の一実施形態による映像信号探索システムにおけるサブ画面の説明図である。

[図15]図15は、本発明の一実施形態による信号探索システムの構成例を示すブロック図である。

[図16]図16は、本発明の一実施例による信号探索システムにおける閾値の設定の説明に用いるグラフである。

[図17]図17は、本発明の一実施例による信号探索システムにおける閾値の設定の説明に用いるグラフである。

[図18]図18は、図15の信号探索システムの動作の一例を示すフローチャートである。

[図19]図19は、本発明の一実施形態による信号探索システムにおける時間窓とフレームの関係の説明図である。

[図20]図20は、本発明の一実施形態による信号探索システムにおけるサブ画面の説明図である。

[図21]図21は、本発明の一実施形態による非線形量子化の説明図である。

[図22]図22は、図15の信号探索システムの動作例の他の例を示すフローチャートである。

WO 2006/006528

14

PCT/JP2005/012667

## 符号の説明

- [0050] 11 目的特徴計算部  
12 蓄積特徴計算部  
13 目的特徴正規化部  
14 蓄積特徴正規化部  
15 目的特徴エリア選択部  
16 蓄積特徴エリア選択部  
17 特徴照合部  
18 蓄積エリア選択特徴データベース  
21 目的特徴計算部  
22 蓄積特徴計算部  
23 目的特徴正規化部  
24 蓄積特徴正規化部  
25 目的特徴量子化部  
26 蓄積特徴量子化部  
27 特徴照合部  
28 蓄積量子化特徴データベース  
31 目的特徴計算部  
32 蓄積特徴計算部  
33 目的特徴正規化部  
34 蓄積特徴正規化部  
35 目的特徴エリア選択部  
36 蓄積特徴エリア選択部  
37 特徴照合部  
38 蓄積エリア選択特徴データベース  
41 目的特徴計算部  
42 蓄積特徴計算部  
43 目的特徴正規化部

WO 2006/006528

15

PCT/JP2005/012667

- 44 蓄積特徴正規化部
- 45 目的特徴量子化部
- 46 蓄積特徴量子化部
- 47 特徴照合部
- 48 蓄積量子化特徴データベース
- 51 目的特徴計算部
- 52 蓄積特徴計算部
- 53 目的特徴正規化部
- 54 蓄積特徴正規化部
- 55 目的特徴エリア選択部
- 56 蓄積特徴エリア選択部
- 57 目的特徴非線形量子化部
- 58 蓄積特徴非線形量子化部
- 59 特徴照合部
- 510 蓄積非線形量子化特徴データベース

#### 発明を実施するための最良の形態

[0051] 以下、図面を参照しつつ、本発明の好適な実施例について説明する。ただし、本発明は以下の各実施例に限定されるものではなく、例えばこれら実施例の構成要素同士を適宜組み合わせてもよい。

[0052] <第1の実施の形態>

[0053] 図1は、本発明における第1の実施の形態を表しており、音響信号を対象とする特徴ひずみに頑健な音響信号検出システムの構成を示すブロック図である。

この図1に示す音響信号検出システムは、音響信号を対象とする特徴ひずみに頑健な信号検出を実現するものであり、目的特徴計算部11と、蓄積特徴計算部12と、目的特徴正規化部13と、蓄積特徴正規化部14と、目的特徴エリア選択部15と、蓄積特徴エリア選択部16と、特徴照合部17と、蓄積エリア選択特徴データベース18とで構成され、蓄積時系列信号(蓄積音響信号)すなわち検索したい音響信号と、目

WO 2006/006528

16

PCT/JP2005/012667

的時系列信号(目的音響信号)すなわち検索される音響信号を入力し、目的時系列信号と類似した蓄積時系列信号中の箇所を出力する。

目的特徴計算部11は、目的音響信号を所定の間隔で時系列にサンプリングして得られる離散値である目的時系列信号から、例えば、サンプリング単位に周波数毎のパワースペクトル値等を抽出し、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる目的特徴を導くものである。

蓄積特徴計算部12も、同様に、蓄積音響信号を所定の間隔で時系列にサンプリングして得られる離散値である蓄積時系列信号から、例えば、サンプリング単位に周波数毎のパワースペクトル値等を抽出し、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる蓄積特徴を導くものである。

[0054] 目的特徴正規化部13は、上記目的特徴から、隣接した部分も含む周辺の目的特徴から導いた統計量を用いて特徴ベクトルの各要素ごと独立に正規化し、正規化された値からなる多次元ベクトルを有する目的正規化特徴を導くものである。

蓄積特徴正規化部14は、上記蓄積特徴から、隣接した部分も含む周辺の蓄積特徴から導いた統計量を用いて特徴ベクトルの各要素ごと独立に正規化し、正規化された値からなる多次元ベクトルを有する蓄積正規化特徴を導くものである。

[0055] 目的特徴エリア選択部15は、上記目的正規化特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、該目的正規化特徴から選択し、該選択された要素の多次元ベクトルからなる目的エリア選択特徴を導くものである。

蓄積特徴エリア選択部16は、上記蓄積正規化特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、該蓄積正規化特徴から選択し、該選択された要素の多次元ベクトルからなる蓄積エリア選択特徴を導くものである。

[0056] 例えば、各エリア選択部は、前記正規化における蓄積特徴および目的特徴の「0」と「1」の境界面との差分を統計量として求め、これを閾値に比較することにより各々エリア特徴を導くことができる。この場合、該閾値は、最も大きい差分の値の7割または8割の数値とすることが考えられる。

各周波数帯において、パワースペクトルの変動の大きい要素を特徴的な要素とするため、上述したように、大きな値を有する要素を特徴パターンとして選択する。

WO 2006/006528

17

PCT/JP2005/012667

このとき、要素毎の下限値を設けておき、この下限値(1つまたは複数個)を超える要素が選択されるまで、閾値を徐々に下げるように制御しても良い。

[0057] また、目的特徴エリア選択部15は、上記目的特徴において、配列における所定の範囲の特徴ベクトルの各要素の平均値(周波数帯域毎の平均値)を求め、各要素からこの平均値を除算した結果の絶対値が所定の閾値を超えた要素を、該目的正規化特徴から選択し、該選択された要素の多次元ベクトルからなる目的エリア選択特徴を導いても良い。

同様に、蓄積特徴エリア選択部16は、上記蓄積特徴において、配列における所定の範囲の特徴ベクトルの各要素の平均値(周波数帯域毎の平均値)を求め、各要素からこの平均値を除算した結果の絶対値が所定の閾値を超えた要素を、該蓄積エリア選択特徴から選択し、該選択された要素の多次元ベクトルからなる蓄積エリア選択特徴を導いても良い。

[0058] 上述の方法により各エリア選択部が、蓄積特徴または目的特徴の平均値との差分の値を統計量として計算し、これを閾値に比較することにより各々エリア選択特徴を導く場合、該閾値は、最も大きい差分の値の7割または8割の数値とすることが考えられる。

各周波数帯において、パワースペクトルの変動の大きい要素を特徴的な要素とするため、上述したように、大きな値を有する要素を特徴パターンとして選択する。

このとき、要素の下限値を設けておき、この下限値を超える要素が選択されるまで、閾値を徐々に下げるように制御しても良い。

[0059] さらに、目的特徴エリア選択部15は、上記目的特徴及び目的正規化特徴を入力し、この目的特徴の配列における所定の範囲(一定区画)の特徴ベクトルの各要素毎の標準偏差(周波数帯域毎の標準偏差)を求め、次に、目的正規化特徴の対応する配列位置の特徴ベクトルの各要素毎に、上記標準偏差を乗算して、乗算結果の絶対値を計算して統計量とし、この統計量において最上位または最上位から複数個(例えば、2個)を、目的正規化特徴から選択して、特徴パターンである、選択された要素の多次元ベクトルからなる目的エリア選択特徴を導いても良い。

[0060] 同様に、蓄積特徴エリア選択部16は、上記蓄積特徴及び蓄積エリア選択特徴を入

WO 2006/006528

18

PCT/JP2005/012667

力し、この蓄積特徴の配列における所定の範囲(一定区画)の特徴ベクトルの各要素毎の標準偏差(周波数帯域毎の標準偏差)を求め、次に、蓄積正規化特徴の対応する配列位置の特徴ベクトルの各要素毎に、上記標準偏差を乗算して、乗算結果の絶対値を計算して統計量とし、この統計量において最上位または最上位から複数個(例えば、2個)を、蓄積正規化特徴から選択して、特徴パターンである、選択された要素の多次元ベクトルからなる蓄積エリア選択特徴を導いても良い。

この目的エリア選択特徴及び蓄積エリア選択特徴は、各々時系列にサンプリングした順序に、各サンプリングにおける特徴ベクトルが順次配列している。配列数(配列の長さ)は蓄積エリア選択特徴より、目的エリア選択特徴の方が短い、または等しい。

[0061] 特徴照合部17は、上記蓄積エリア選択特徴における時系列に配列している特徴ベクトルにおいて、この配列の所定の範囲を照合区間として設定し、この照合区間と同一の長さの被照合区間を目的エリア選択特徴に設定して、上記照合区間と被照合区間との類似度を計算し、予め設定された探索閾値と比較して、類似しているか否かの判定を行う。

また、特徴照合部17は、上記照合区間の照合処理が終了すると、新たな照合区間を設定するため、配列の隣接する同一時間幅の範囲にずらす。

蓄積エリア選択特徴データベース18は、複数の、例えば良く配信される曲に対して、蓄積特徴計算部12、蓄積特徴正規化部14及び蓄積特徴エリア選択部16により、予め蓄積エリア選択特徴を計算して、各曲名に対応して記憶している。

[0062] 次に、図面を参照して本実施例の音響信号検出システムの動作の説明を行う。図2は図1の音響信号検出システムの動作例を示すフローチャートである。

蓄積特徴計算部12は、与えられた蓄積音響信号を読み込んで入力し(ステップS1)、入力した蓄積音響信号に対して特徴抽出を行う。

このとき、蓄積特徴計算部12は、音響信号のフーリエ変換の振幅成分を用い、(例えば周波数8000Hzで標本化した音響信号の1秒の区間をフーリエ変換し、0~4000Hzを等間隔に32の周波数帯域の区間に分割し、0.1秒ごとに配列する)、各区間内での振幅成分の平均パワーからなる32次元の多次元ベクトルを特徴ベクトルとして抽出して蓄積特徴とする(ステップS2)。

WO 2006/006528

19

PCT/JP2005/012667

[0063] 次に、蓄積特徴正規化部14は、蓄積特徴計算部12から蓄積特徴を読み込み、この蓄積特徴の特徴ベクトルの各要素毎に、所定の一定区画の平均値と標準偏差とを計算して求める。

例えば、蓄積特徴正規化部14は、各周波数帯域(要素)ごとにある時間区間の値から平均値と標準偏差とを求め、求められた平均値と標準偏差とを用いて正規化を行う。

このとき、蓄積特徴正規化部14による正規化後の周波数特徴 $y(i)$ の $k$ 番目の要素は、以下の(1)式となる。

[0064] [数1]

$$y(i,k) = \frac{1}{\sigma(i,k)} (x(i,k) - m(i,k)) \quad \dots(1)$$

[0065] ただし、(1)式において、 $m(i,k)$ は平均値であり、以下に示す(2)式により求められ、 $\sigma(i,k)$ は標準偏差であり、以下に示す(3)式により求められる。

[0066] [数2]

$$m(i,k) = \frac{1}{2M} \sum_{i=-M}^{M-1} x(i,k) \quad \dots(2)$$

[0067] [数3]

$$\sigma(i,k)^2 = \frac{1}{2M} \sum_{i=-M}^{M-1} (x(i,k) - m(i,k))^2 \quad \dots(3)$$

[0068] (2)式及び(3)式において、 $M$ は局所時間内の周波数特徴の平均値及び標準偏

WO 2006/006528

20

PCT/JP2005/012667

差を求めるための時間窓の大きさの半分の数値である。

そして、(1)～(3)式により得られた多次元ベクトルの時系列な配列(kで示される順番)を蓄積正規化特徴とする(ステップS3)。

[0069] 次に、蓄積特徴エリア選択部16は、蓄積特徴計算部12から蓄積特徴を、また、蓄積特徴正規化部14から、蓄積正規化特徴を読み込んで入力し、蓄積特徴の多次元ベクトルの各要素毎の標準偏差 $\sigma(i,k)$ を、各要素(周波数帯域のパワースペクトルからなる特徴ベクトル)の配列の一定区画において、上記(2)及び(3)式を用いて求める。

そして、蓄積特徴エリア選択部16は、以下の(5)式に示すように、各要素 $y(i,k)$ 毎に上記標準偏差 $\sigma(i,k)$ を乗じて、その絶対値をとる。

[0070] [数4]

$$z(i,k) = | y(i,k) \cdot \sigma(i,k) | \quad \dots(4)$$

[0071] ここで、蓄積特徴エリア選択部16は、得られた各要素と標準偏差との乗算結果の統計量 $z(i,k)$ から、一定区画単位に要素毎に最上位から複数個のベクトル、例えば上位2つのベクトルを選択する。

そして、蓄積特徴エリア選択部16は、上記一定区画において、蓄積正規化特徴から各要素毎に選択された要素のベクトルからなる多次元ベクトルを蓄積エリア選択特徴として出力する(ステップS4)。

[0072] このとき、蓄積特徴エリア選択部16は、計算により得られた蓄積エリア選択特徴を、特徴照合部17へ直接送信するか、または一旦蓄積エリア選択特徴データベース18への登録のいずれかの処理を行う。

特徴照合部17がリアルタイムに蓄積エリア選択特徴と、目的エリア選択特徴との比較を行う場合、蓄積特徴エリア選択部16は特徴照合部17へ、入力されている蓄積音響信号の蓄積エリア選択特徴を出力し、蓄積エリア選択特徴データベース18へ蓄積音響信号のデータを登録する場合、蓄積特徴エリア選択部16は特徴照合部17



WO 2006/006528

21

PCT/JP2005/012667

へ蓄積エリア選択特徴を送信せずに、蓄積エリア選択特徴データベース18へ、曲名に対応させて蓄積エリア選択特徴の登録処理を行う。

[0073] 目的特徴計算部11は、与えられた目的音響信号を読み込んで入力し(ステップS5)、入力した目的音響信号に対して特徴抽出を行う。

このとき、目的特徴計算部11は、蓄積特徴計算部12と同様に、音響信号のフーリエ変換の振幅成分を用い、(例えば周波数8000Hzで標本化した音響信号の1秒の区間をフーリエ変換し、0～4000Hzを等間隔に32の周波数帯域の区間(特徴ベクトルの要素)に分割し、0.1秒ごとに配列する)、各区間内での振幅成分の平均パワーからなる32次元の多次元ベクトルを特徴ベクトルとして抽出して目的特徴とする(ステップS6)。

[0074] 次に、目的特徴正規化部13は、目的特徴計算部11から目的特徴を読み込み、この目的特徴の特徴ベクトルの各要素毎に、所定の一定区画の平均値と標準偏差とを計算して求める。

すなわち、目的特徴正規化部13は、蓄積特徴正規化部14と同様に、(1)～(3)式により得られた多次元ベクトルの時系列な配列(kで示される順番)を目的正規化特徴とする(ステップS7)。

[0075] 次に、目的特徴エリア選択部15は、目的特徴計算部11から目的特徴を、また目的特徴正規化部13から目的正規化特徴を読み込んで入力し、蓄積特徴エリア選択部16と同様に、目的特徴の一定区画から求めた標準偏差 $\sigma(i,k)$ を、各要素毎に乗算して乗算結果を統計量とする。

そして、目的特徴エリア選択部15は、各要素毎の配列において、上記統計量から数値の大きい順に上位2つの要素を選択して、目的正規化特徴から各要素毎に選択された要素のベクトルからなる多次元ベクトルを目的エリア選択特徴として出力する(ステップS8)。

[0076] 次に、特徴照合部17は、目的特徴エリア選択部15及び蓄積特徴エリア選択部16により得られた目的エリア選択特徴及び蓄積エリア選択特徴を各々読み込む。

また、特徴照合部17は、目的音響信号及び蓄積音響信号を同時に入力して、リアルタイムに類似を判定する場合以外、蓄積エリア選択特徴データベース18から順次

WO 2006/006528

22

PCT/JP2005/012667

比較を行う蓄積エリア選択特徴を読み出して、目的エリア選択特徴と比較する。

このとき、特徴照合部17は、蓄積エリア選択特徴において、目的特徴エリア選択部15で与えられた目的エリア選択特徴と同じ長さの特徴ベクトルの配列を照合区間として設定する。

すなわち、特徴照合部17は、複数の上記一定区画からなる目的エリア選択特徴を照合区間とし、蓄積エリア選択特徴において、目的エリア選択特徴と同一の一定区画数からなる配列長を、目的エリア選択特徴に対する照合区間として対応させて、順次、両照合区間を類似度により比較する。

[0077] そして、特徴照合部17は、目的エリア選択特徴と蓄積エリア選択特徴とにおける照合区間の比較箇所との類似度の計算を、下記(5)式を用いて行う。

すなわち、特徴照合部17は、目的エリア選択特徴における要素 $y_q(i, k)$ 及び蓄積エリア選択特徴の照合区間における要素 $y_s(i, k)$ の一致箇所のユークリッド距離と、非一致箇所の数に所定の定数 $a$ を乗じた値の和を計算して、これを類似度として出力する(ステップS9)。

[0078] [数5]

$$S = \sum_{k=q \cap s} (y_q(i, k) - y_s(i + j, k))^2 + a(\max(|q|, |s|) - |q \cap s|) \quad \dots(5)$$

[0079] ただし、 $q, s$ は目的特徴及び蓄積特徴のエリア選択箇所、 $x \cap y$ は $x$ と $y$ の共通箇所、 $|x|$ はエリア選択箇所 $x$ の要素数、 $\max(x, y)$ は $x$ と $y$ のうち大きい方の数を表す。

ここで、例えば、特徴照合部17は、目的エリア選択特徴が15秒の長さであるとする、この目的エリア選択特徴の特徴ベクトルの配列から、配列の要素としての特徴ベクトルを、0.1秒間隔に合計150箇所抽出し、各サンプリングにおいて周波数帯域を32分割しているため、これらのベクトルからなる $150 \times 32$ の4800次元から、各周波数帯域毎にパワースペクトルの上位2つ(目的特徴エリア選択部15が抽出)を用い、64の要素からなる多次元ベクトルを照合に用いる目的ベクトルとする。

[0080] また、特徴照合部17は、上記目的ベクトルと同様に、蓄積エリア選択特徴のデータ

WO 2006/006528

23

PCT/JP2005/012667

の先頭から、15秒の長さを単位に順次照合区間として設定して、特徴ベクトルの配列から、特徴ベクトルを0.1秒間隔に合計150箇所抽出し、各サンプリングにおいて周波数帯域を32分割しているため、それらからなる $150 \times 32$ の4800次元から、各周波数帯域毎にパワースペクトルの上位2つ(蓄積特徴エリア選択部16が抽出)を用い、64の要素からなる多次元ベクトルを照合に用いる蓄積ベクトルとする。

- [0081] そして、特徴照合部17は、照合区間を、蓄積エリア選択特徴の先頭から順次ずらしながら、目的エリア選択特徴との、すなわち、(5)式を用いて上記目的ベクトルと蓄積ベクトルとの類似度を計算し、予め設定した探索閾値と比較処理を行う(ステップS10)。

目的エリア選択特徴と蓄積エリア選択特徴とを、蓄積エリア選択特徴の最後まで照合処理した後、照合区間毎の類似度が予め設定された探索閾値より低い、蓄積エリア選択特徴の照合区間の領域を探索結果として出力する(ステップS11)。

なお、特徴照合部17は、探索結果として、比較した結果において、もっとも類似度が低い蓄積エリア選択特徴の照合区間の領域を出力するようにしてもよい。

- [0082] また、特徴照合部17は、探索閾値を複数の照合区間の類似度が下回った場合、類似度の上位(低い数値のものから)N番目までの照合区間を出力するようにすることも可能である。

さらに、特徴照合部17は、探索閾値を下回る照合区間が無い場合など、該当箇所がないことを示す情報を通知し、新たな蓄積音響信号の蓄積エリア選択特徴を、蓄積エリア選択特徴データベース18から読み出し、上記探索閾値以下の照合区間を有する蓄積エリア選択特徴が探索されるまで、ステップS9以降の探索処理を継続させて行うようにしてもよい。

- [0083] また、図1における目的特徴計算部11、目的特徴正規化部13及び目的特徴エリア選択部5を各ユーザの端末(例えば、パーソナルコンピュータ)にインストールしておき、音楽配信を行うサービスプロバイダに蓄積特徴計算部12、蓄積特徴正規化部14、蓄積特徴エリア選択部16、特徴照合部17及び蓄積エリア選択特徴データベース18を有する音響信号検出サーバを設けてもよい。

これにより、ユーザが携帯電話等で受信した音響信号を目的音響信号として、目的

WO 2006/006528

24

PCT/JP2005/012667

エリア選択特徴まで生成し、この目的エリア選択特徴を上記音響信号検出サーバへ、インターネットなどを介して送信し、この目的エリア選択特徴に類似した蓄積音響信号を探索してもらうように要求する構成とする。このとき蓄積エリア選択特徴及び目的エリア選択特徴を計算するときの一定区画の特徴ベクトルの配列長等の規定は、音響信号検出サーバと端末とにおいて予め一致させておく。

[0084] 次に、上記音響信号検出システムを適用した場合の動作実験例及びその結果を示す。

本発明の音響信号検出システムによる効果を確認するため、本発明を適用した場合と適用しなかった場合との探索精度の比較を行った。

実験のため、蓄積音響信号として、あるCD(コンパクトディスク)の曲を再生した音響信号を、そのまま図1の音響信号検出システムの蓄積特徴計算部12に入力させた。

一方、目的音響信号として、雑音の大きい喫茶店内のスピーカにより、上記CDの同一の曲の所定の部分(蓄積音響信号の一部分)を鳴らし、その喫茶店内で携帯電話をPHS(登録商標)にかけ、そのPHS(登録商標)の音声出力から受音して得た音響信号を、目的音響信号として、目的特徴計算部11へ入力させた。

[0085] 上述した同一の実験条件において、200回繰り返して探索処理を行い、探索精度の測定を行った。

この探索精度は、探索閾値の設定を調節して、適合率(precision rate)と再現率(recall rate)とが等しくなるときの値とした。

ここで、適合率とは探索結果として出力されたもののうち正しいものの割合であり、再現率とは探索されるべきもののうち探索結果として出力されたものの割合である。

適合率または再現率は、探索閾値の設定によって変化するが、本実験においては探索閾値((5)式で用いる類似度に対する閾値)を以下に示す(6)式によって定めた。

[0086] [数6]

$$\theta = m + n\nu \quad \dots(6)$$

WO 2006/006528

25

PCT/JP2005/012667

[0087] 上記(6)式において、 $m$ 及び $v$ 各々は、与えられた蓄積信号に対して入力信号をサンプリングし、予備的に類似度の計算を行って収集した類似度の平均値と標準偏差であり、 $n$ は経験的に与えられる係数である。

ただし、(6)式において、探索閾値  $\theta$  が1を超える場合、 $\theta = 1$ とし、0を下回るときは  $\theta = 0$ とした。

本実験においては、(6)式における $n$ の値を200回繰り返して行う間で一定とし、その $t$ の値を適合率と再現率とがほぼ等しくなるように調節した。

上述した実験の結果、上記精度は、正規化特徴により得られた蓄積及び目的エリア選択特徴を用いない場合に15.0%であり、本発明における蓄積及び目的エリア選択特徴を用いた場合(一実施例)で80.1%であった。

[0088] また、探索処理においては、CPUにIntel(登録商標)Pentium(登録商標) III 1133MHz、OSにRedHat(登録商標)7.3、コンパイラにGNU gccを用いた。

また、実行ファイルはコンパイラ最適化オプション「-O3」によりコンパイルを行った。

上述した実験において、本発明の音響信号検出システムにより、探索精度が向上したことを確認することができる。

[0089] 本発明の音響信号検出システムは、実環境で収録した雑音やとぎれのある断片的な音響信号をつかって、一致する音楽やCMの曲名や放送時間を探索するだけでなく、音楽やCMに関連する情報を蓄積した付随情報データベースと組み合わせて情報探索するのにも利用できる。

例えば、TVやラジオから流れる音楽或いはCMをユーザが携帯電話で受信し、音響信号検索サービスなどに送信する。音響信号検索サービス提供者は、この音響信号に一致又は類似する音響をデータベースより検索し、音楽或いはこのCMに関する情報(例えば演奏者、作詞作曲者名、商品情報、製品特徴、サービス、撮影地、出演者、ホームページ等)をインターネットなどのネットワークを介して有料、または無料でユーザに提供するといった構成が可能である。

[0090] ここで、検索する音響の入力方法においては、携帯端末からサーバにアクセスし、

WO 2006/006528

26

PCT/JP2005/012667

端末のマイクで直接信号を受信し目的信号としてもかまわない。また、いったん録音した音響を適当に切り取って送信してもよい。

また、他の発明の実施例として、ユーザが家庭用テープレコーダで録音したイメージ音響から、そのイメージに近い音楽のフレーズを配信されているCDや映画、テレビプログラムから検索し、コンテンツを利用したり、また録音したイメージ音響に相当するフレーズを集めて編集したりすることにも利用できるであろう。

[0091] <第2の実施の形態>

図3は、本発明の第2の実施の形態を表しており、音響信号を対象とする特徴ひずみに頑健な音響信号検出システムの構成を示すブロック図である。

この図3に示す音響信号検出システムは、音響信号を対象とする特徴ひずみに頑健な信号検出を実現するものであり、目的特徴計算部21と、蓄積特徴計算部22と、目的特徴正規化部23と、蓄積特徴正規化部24と、目的特徴量子化部25と、蓄積特徴量子化部26と、特徴照合部27と、蓄積量子化特徴データベース28とで構成され、蓄積時系列信号(蓄積音響信号)すなわち検索したい音響信号と、目的時系列信号(目的音響信号)すなわち検索される音響信号を入力し、目的時系列信号と類似した蓄積時系列信号中の箇所を出力する。

[0092] 目的特徴計算部21は、目的音響信号を所定の間隔で時系列にサンプリングして得られる離散値である目的時系列信号から、例えば、サンプリング単位に周波数毎のパワースペクトル値等を抽出し、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる目的特徴を導く(計算する)ものである。

蓄積特徴計算部22も、同様に、蓄積音響信号を所定の間隔で時系列にサンプリングして得られる離散値である蓄積時系列信号から、例えば、サンプリング単位に周波数毎のパワースペクトル値等を抽出し、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる蓄積特徴を導くものである。

[0093] 目的特徴正規化部23は、上記目的特徴から、隣接した部分も含む周辺の目的特徴から導いた統計量を用いて特徴ベクトルの各要素ごと独立に正規化し、正規化された値からなる多次元ベクトルを有する目的正規化特徴を導くものである。

蓄積特徴正規化部24は、上記蓄積特徴から、隣接した部分も含む周辺の蓄積特

WO 2006/006528

27

PCT/JP2005/012667

徴から導いた統計量を用いて特徴ベクトルの各要素ごと独立に正規化し、正規化された値からなる多次元ベクトルを有する蓄積正規化特徴を導くものである。

[0094] 目的特徴量子化部25は、上記目的正規化特徴を所定の閾値によりスカラー量子化し、この量子化により得られた値を要素とする多次元ベクトルからなる目的量子化特徴を導くものである。

蓄積特徴量子化部26は、上記蓄積正規化特徴を所定の閾値によりスカラー量子化し、この量子化により得られた値を要素とする多次元ベクトルからなる蓄積量子化特徴を導くものである。

[0095] 量子化を行う上記所定の閾値は、目的及び蓄積正規化特徴を、例えば2値化する場合、2値化して得られる要素の数値と、もとの要素の数値との平均2乗誤差の最小化される点を選択する等で求められる。

この目的量子化特徴及び蓄積量子化特徴は、各々時系列にサンプリングした順序に、各サンプリングにおける特徴ベクトルが順次配列している。配列数(配列の長さ)は蓄積量子化特徴より、目的量子化特徴の方が短いまたは等しい。

なお、以降において、蓄積特徴についての縮退させたベクトルとは、蓄積エリア選択特徴や蓄積正規化特徴等の所定の統計量に基づいて計算により得られるベクトルであり、また、目的特徴についての縮退させたベクトルとは、目的エリア選択特徴や目的正規化特徴等の所定の統計量に基づいて計算により得られるベクトルである。

[0096] 特徴照合部27は、上記蓄積量子化特徴における時系列に配列している特徴ベクトルにおいて、この配列の所定の範囲を照合区間として設定し、この照合区間と同一の長さの被照合区間を目的量子化特徴に設定して、上記照合区間と被照合区間との類似度を計算し、予め設定された探索閾値と比較して、類似しているか否かの判定を行う。

また、特徴照合部27は、上記照合区間の照合処理が終了すると、新たな照合区間を設定するため、配列の隣接する同一時間幅の範囲にずらす。

蓄積量子化特徴データベース28は、複数の、例えば良く配信される曲に対して、蓄積特徴計算部22、蓄積特徴正規化部24及び蓄積特徴量子化部26により、予め蓄積量子化特徴を計算して、各曲名に対応して記憶されている。

WO 2006/006528

28

PCT/JP2005/012667

[0097] 次に、図面を参照して本実施例の音響信号検出システムの動作の説明を行う。図4は図3の音響信号検出システムの動作例を示すフローチャートである。

蓄積特徴計算部22は、与えられた蓄積音響信号を読み込んで入力し(ステップT1)、入力した蓄積音響信号に対して特徴抽出を行う。

このとき、蓄積特徴計算部22は、音響信号のフーリエ変換の振幅成分を用い、例えば周波数8000Hzで標本化した音響信号の1秒の区間をフーリエ変換し、0～4000Hzを等間隔に32の周波数帯域の区間に分割し、0.1秒ごとに、各区間内での振幅成分の平均パワーからなる32次元の多次元ベクトルを特徴ベクトルとして抽出して蓄積特徴とする(ステップT2)。

[0098] 次に、蓄積特徴正規化部24は、蓄積特徴計算部22から蓄積特徴を読み込み、この蓄積特徴の特徴ベクトルの各要素毎に、所定の一定区画の平均値と標準偏差とを計算して求める。

例えば、蓄積特徴正規化部24は、各周波数帯域ごとにある時間区間の値から平均値と標準偏差とを求め、求められた平均値と標準偏差とを用いて正規化を行う。

このとき、蓄積特徴正規化部24による正規化後の周波数特徴 $y(i)$ の $k$ 番目の要素は、以下の(21)式となる。

[0099] [数21]

$$y(i,k) = \frac{1}{\sigma(i,k)} (Q(i,k) - m(i,k)) \quad \dots (21)$$

[0100] ただし、(21)式において、 $m(i,k)$ は平均値であり、以下に示す(22)式により求められ、 $\sigma(i,k)$ は標準偏差であり、以下に示す(23)式により求められる。

[0101] [数22]

$$m(i,k) = \frac{1}{2M} \sum_{l=-M}^{M-1} Q(i,k) \quad \dots (22)$$



WO 2006/006528

29

PCT/JP2005/012667

[0102] [数23]

$$\sigma(i,k)^2 = \frac{1}{2M} \sum_{i=-M}^{M-1} (Q(i,k) - m(i,k))^2 \quad \dots (23)$$

[0103] (22)式及び(23)式において、Mは局所時間内の周波数特徴の平均値及び標準偏差を求めるための時間窓の大きさの半分の数値である。

そして、(21)～(23)式により得られた多次元ベクトルの時系列な配列(kで示される順番)を蓄積正規化特徴とする(ステップT3)。

[0104] 次に、蓄積特徴量子化部26は、蓄積特徴正規化部24から、蓄積正規化特徴を読み込んで入力し、蓄積正規化特徴の多次元ベクトルの各要素毎に、あらかじめ与えられた閾値を境に二値化処理をして量子化された数値のベクトルを得る。

例えば、蓄積特徴量子化部26は、蓄積正規化特徴の各要素を、得られた閾値tを境に量子化する。

これにより、蓄積特徴量子化部26は、量子化後の周波数特徴z(i)のk番目の要素を、以下に示す(24)式により各々計算する。

[0105] [数24]

$$z(i,k) = \begin{cases} 1 & (y(i,k) > t \text{ のとき}) \\ 0 & (y(i,k) \leq t \text{ のとき}) \end{cases} \quad \dots (24)$$

[0106] そして、蓄積特徴量子化部26は、各要素を2値化して、この多次元ベクトルを蓄積量子化特徴として出力する(ステップT4)。

このとき、蓄積特徴量子化部26は、計算により得られた蓄積量子化特徴を、特徴照合部27へ直接送信するか、または一旦蓄積量子化特徴データベース28への登録のいずれかの処理を行う。

特徴照合部27がリアルタイムに蓄積量子化特徴と、目的量子化特徴との比較を行う場合、蓄積特徴量子化部26は特徴照合部27へ、入力されている蓄積音響信号の

WO 2006/006528

30

PCT/JP2005/012667

蓄積量子化特徴を出力し、蓄積量子化特徴データベース28へ蓄積音響信号のデータを登録する場合、蓄積特徴量子化部26は特徴照合部27へ蓄積量子化特徴を送信せずに、蓄積量子化特徴データベース28へ、曲名に対応させて蓄積量子化特徴の登録処理を行う。

- [0107] 目的特徴計算部21は、与えられた目的音響信号を読み込んで入力し(ステップT5)、入力した目的音響信号に対して特徴抽出を行う。

このとき、目的特徴計算部21は、蓄積特徴計算部22と同様に、音響信号のフーリエ変換の振幅成分を用い、例えば周波数8000Hzで標本化した音響信号の1秒の区間をフーリエ変換し、0～4000Hzを等間隔に32の周波数帯域の区間に分割し、0.1秒ごとに、各区間内での振幅成分の平均パワーからなる32次元の多次元ベクトルを特徴ベクトルとして抽出して目的特徴とする(ステップT6)。

- [0108] 次に、目的特徴正規化部23は、目的特徴計算部21から目的特徴を読み込み、この目的特徴の特徴ベクトルの各要素毎に、所定の一定区画の平均値と標準偏差とを計算して求める。

すなわち、目的特徴正規化部23は、蓄積特徴正規化部24と同様に、(21)～(23)式により得られた多次元ベクトルの時系列な配列(kで示される順番)を目的正規化特徴とする(ステップT7)。

- [0109] 次に、目的特徴量子化部25は、目的特徴正規化部23から、目的正規化特徴を読み込んで入力し、目的正規化特徴の多次元ベクトルの各要素毎に、あらかじめ与えられた閾値を境に、蓄積特徴量子化部26と同様に、(24)式を用いて二値化処理をして、量子化された数値のベクトルを得る。

そして、目的特徴量子化部25は、各要素が2値化された、この多次元ベクトルを目的量子化特徴として出力する(ステップT8)。

- [0110] 次に、特徴照合部27は、目的特徴量子化部25及び蓄積特徴量子化部26により得られた目的量子化特徴及び蓄積量子化特徴を各々読み込む。

また、特徴照合部27は、目的音響信号及び蓄積音響信号を同時に入力して、リアルタイムに類似を判定する場合以外、蓄積量子化特徴データベース28から順次比較を行う蓄積量子化特徴を読み出して、目的量子化特徴と比較する。

WO 2006/006528

31

PCT/JP2005/012667

このとき、特徴照合部27は、蓄積量子化特徴において、目的特徴量子化部25で与えられた目的量子化特徴と同じ長さの特徴ベクトルの配列を照合区間として設定する。

[0111] そして、特徴照合部27は、目的量子化特徴と、上記照合区間との類似度の計算を、類似度として、双方の特徴ベクトル間のハミング距離を計算して行う(ステップT9)。

例えば、特徴照合部27は、目的量子化特徴が15秒の長さであるとする、この目的量子化特徴の特徴ベクトルの配列から、配列の要素としての特徴ベクトルを、0.1秒間隔に合計150箇所抽出し、各サンプリングにおいて周波数帯域を32分割しているため、これらのベクトルからなる $150 \times 32$ の4800次元を照合に用いる目的ベクトルとする。

[0112] また、特徴照合部27は、上記目的ベクトルと同様に、蓄積量子化特徴のデータの先頭から、15秒の長さを単位に順次照合区間として設定して、特徴ベクトルの配列から、特徴ベクトルを0.1秒間隔に合計150箇所抽出し、各サンプリングにおいて周波数帯域を32分割しているため、それらからなる $150 \times 32$ の4800次元を照合に用いる蓄積ベクトルとする。

[0113] このとき、特徴照合部27が目的量子化特徴の特徴ベクトルの配列から、複数の箇所の要素を抽出して、目的ベクトルとするのであれば、目的特徴量子化部25において、予め配列の要素としての特徴ベクトルを目的正規化特徴から抽出して、すなわち、0.1秒間隔に合計150箇所抽出して、スカラー量子化を行い、目的ベクトルとして特徴照合部27へ出力するようにしてもよい。

[0114] このとき、特徴照合部27は、照合区間を、蓄積量子化特徴の先頭から順次ずらしながら、目的量子化特徴との、すなわち上記目的ベクトルと蓄積ベクトルとのハミング距離を計算することにより比較処理を行う。

そして、特徴照合部27は、蓄積量子化特徴を最後まで照合処理した後、照合区間毎のハミング距離を参照して、ハミング距離が最も小さい照合区間の領域を探索結果として出力する。

なお、特徴照合部27は、探索結果として、予めハミング距離の探索閾値を与えられていた場合、この探索閾値と選択された照合区間のハミング距離とを判定して(ステッ

WO 2006/006528

32

PCT/JP2005/012667

プT10)、この探索閾値を下回るもののみ出力することもできる(ステップT11)。

- [0115] また、特徴照合部27は、探索閾値を複数の照合区間のハミング距離が下回った場合、ハミング距離の上位(低い数値のものから)N番目までの照合区間を出力するようにすることも可能である。

さらに、特徴照合部27は、探索閾値を下回る照合区間が無い場合など、該当箇所がないことを示す情報を通知し、新たな蓄積音響信号の蓄積量子化特徴を、蓄積量子化特徴データベース28から読み出し、上記探索閾値以下の照合区間を有する蓄積量子化特徴が探索されるまで、ステップT9以降の探索処理を継続させて行うようにしてもよい。

- [0116] また、図3における目的特徴計算部21、目的特徴正規化部23及び目的特徴量子化部25を各ユーザの端末(例えば、パーソナルコンピュータ)にインストールしておき、音楽配信を行うサービスプロバイダに蓄積特徴計算部22、蓄積特徴正規化部24、蓄積特徴量子化部26、特徴照合部27及び蓄積量子化特徴データベース28を有する音響信号検出サーバを設けてもよい。

これにより、ユーザが携帯電話等で受信した音響信号を目的音響信号として、目的量子化特徴まで生成し、この目的量子化特徴を上記音響信号検出サーバへ、インターネットなどを介して送信し、この目的量子化特徴に類似した蓄積音響信号を探索してもらうように要求する構成とする。

- [0117] 次に、上記音響信号検出システムを適用した場合の動作実験例及びその結果を示す。

本発明の音響信号検出システムによる効果を確認するため、本発明を適用した場合と適用しなかった場合との探索精度の比較を行った。

実験のため、蓄積音響信号として、あるCD(コンパクトディスク)の曲を再生した音響信号を、そのまま図3の音響信号検出システムの蓄積特徴計算部22に入力させた。

一方、目的音響信号として、スピーカーにより、上記CDの同一の曲の所定の部分(蓄積音響信号の一部)を鳴らし、その喫茶店内で携帯電話をPHS(登録商標)にかけ、そのPHS(登録商標)の音声出力から受音して得た音響信号を、目的音響信

WO 2006/006528

33

PCT/JP2005/012667

号として、目的特徴計算部21へ入力させた。

[0118] 上述した同一の実験条件において、200回繰り返して探索処理を行い、探索精度の測定を行った。

この探索精度は、探索閾値の設定を調節して、適合率(precision rate)と再現率(recall rate)とが等しくなるときの値とした。

ここで、適合率とは探索結果として出力されたもののうち正しいものの割合であり、再現率とは探索されるべきもののうち探索結果として出力されたものの割合である。

適合率または再現率は、探索閾値の設定によって変化するが、本実験においては探索閾値(ハミング距離に対する閾値)を以下に示す(25)式によって定めた。

[0119] [数25]

$$\theta = m + n \nu \quad \dots (25)$$

[0120] 上記(25)式において、 $m$ 及び $\nu$ 各々は、与えられた蓄積信号に対して入力信号をサンプリングし、予備的に類似度の計算を行って収集した類似度の平均値と標準偏差であり、 $n$ は経験的に与えられる係数である。

ただし、(25)式において、探索閾値 $\theta$ が1を越える場合、 $\theta = 1$ とし、0を下回るときは $\theta = 0$ とした。

本実験においては、(25)式における $n$ の値を200回繰り返して行う間で一定とし、その $t$ の値を適合率と再現率とがほぼ等しくなるように調節した。

上述した実験の結果、上記精度は、正規化特徴の量子化を行わなかった場合に60.0%であり、量子化を行った場合(一実施例)で85.77%であった。

[0121] また、探索処理においては、CPUにIntel(登録商標)Pentium(登録商標) III 1133MHz、OSにRedHat(登録商標)7.3、コンパイラにGNU gccを用いた。

また、実行ファイルはコンパイラ最適化オプション「-O3」によりコンパイルを行った。

上述した実験において、本発明の音響信号検出システムにより、探索精度が向上

WO 2006/006528

34

PCT/JP2005/012667

したことを確認することができる。

- [0122] 本発明の音響信号検出システムは、実環境で収録した端末特性、符号化特性などによりひずみをうけた断片的な音響信号をつかって、一致する音楽やCMの曲名や放送時間を探索するだけでなく、音楽やCMに関連する情報を蓄積した付随情報データベースと組み合わせて情報探索するのにも利用できる。

例えば、TVやラジオから流れる音楽或いはCMをユーザが携帯電話で受信し、音響信号検索サービスなどに送信する。音響信号検索サービス提供者は、この音響信号に一致又は類似する音響をデータベースより検索し、音楽或いはこのCMに関する情報(例えば演奏者、作詞作曲者名、商品情報、製品特徴、サービス、撮影地、出演者、ホームページ等)をインターネットなどのネットワークを介して有料、または無料でユーザに提供するといった構成が可能である。

- [0123] ここで、検索する音響の入力方法においては、携帯端末からサーバにアクセスし、端末のマイクで直接信号を受信し目的信号としてもかまわない。また、いったん録音した音響を適当に切り取って送信してもよい。

また、他の発明の実施例として、ユーザが家庭用テープレコーダで録音したイメージ音響から、そのイメージに近い音楽のフレーズを配信されているCDや映画、テレビプログラムから検索し、コンテンツを利用したり、また録音したイメージ音響に相当するフレーズを集めて編集したりすることにも利用できるであろう。

- [0124] <第3の実施の形態>

以下、本発明の実施の第3の実施の形態について図面を参照しながら説明する。図5は、本発明の一実施形態であり、映像信号を対象とする反射や途切れに頑健な映像信号探索システムの構成を示すブロック図である。

この図5に示す映像信号探索システムは、映像信号を対象とする反射や途切れに頑健な信号検出を実現するものであり、目的特徴計算部31と、蓄積特徴計算部32と、目的特徴正規化部33と、蓄積特徴正規化部34と、目的特徴エリア選択部35と、蓄積特徴エリア選択部36と、特徴照合部37と、蓄積エリア選択特徴データベース38とで構成されており、蓄積映像信号、即ち検索される映像信号と、目的映像信号、即ち検索したい映像信号を入力し、目的時系列信号に類似した蓄積時系列信号中の箇

WO 2006/006528

35

PCT/JP2005/012667

所を出力する。

[0125] 目的特徴計算部31は、目的映像信号を入力とし、この目的映像信号を所定の間隔で時系列にサンプリングして得られる離散値である目的時系列信号から、例えば、目的映像信号からサンプリング単位に周波数毎のパワースペクトル等を抽出して特徴量とし、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる目的特徴を算出するものである。

[0126] 蓄積特徴計算部32は、目的特徴計算部31と同様に、蓄積映像信号を所定の間隔で時系列にサンプリングして得られる離散データである蓄積時系列信号から、抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる蓄積特徴を算出するものである。

なお、ここでは、目的特徴及び蓄積特徴を、多次元ベクトルとしているが、行列表現するようにしても良い。

[0127] 目的特徴正規化部33は、上記目的特徴から、その目的特徴に隣接した目的特徴を含む周辺の複数の目的特徴から導いた統計量を用いて、特徴ベクトルの各要素毎独立に正規化し、正規化された値からなる特徴ベクトルを有する目的正規化特徴を導くものである。

[0128] 蓄積特徴正規化部34は、上記蓄積特徴から、その蓄積特徴に隣接した蓄積特徴を含む周辺の複数の蓄積特徴から導いた統計量を用いて特徴ベクトルの要素毎、独立に正規化し、正規化された値からなる特徴ベクトルを有する蓄積正規化特徴を導くものである。

ここで、「蓄積特徴に『隣接した』蓄積特徴」の表現で用いられる「隣接した」とは、例えば、映像信号を時間に関してサンプリングした際に得られる離散時刻のフレームに対し、その前後の離散時刻のフレームを表した場合の「前後の離散時刻の」に対応する。また、「周辺の複数の」とは、前記例を用いると、「ある離散時刻間の」に対応する。

[0129] 目的特徴エリア選択部35は、上記目的特徴から所定の統計量を計算し、その統計量が所定の閾値を越える要素を、目的正規化特徴から選択し、選択された要素の多次元ベクトルからなる目的エリア選択特徴を導くものである。

WO 2006/006528

36

PCT/JP2005/012667

[0130] 蓄積特徴エリア選択部36は、上記蓄積正規化特徴を所定の統計量を計算し、その統計量が所定の閾値を越える要素を、蓄積正規化特徴から選択し、選択された要素の多次元ベクトルからなる蓄積エリア選択特徴を導くものである。

[0131] 特徴照合部37は、上記蓄積エリア選択特徴に所定の範囲を照合区間として設定し、この照合区間と同一の長さの被照合区間を目的エリア選択特徴に設定して、上記照合区間と被照合区間との類似度を計算し、予め設定された探索閾値と比較して、類似しているか否かの判定を行い、その結果を信号検出結果として出力する。

また、特徴照合部37は、上記照合区間の照合処理が終了すると、新たな照合区間を設定するため、配列の隣接する同一時間幅の範囲に照合区間をずらす処理を行う。

[0132] 蓄積エリア選択特徴データベース38は、蓄積特徴計算部32、蓄積特徴正規化部34及び蓄積特徴エリア選択部36により予め計算された蓄積エリア選択特徴を記憶している。

複数のCM映像を蓄積映像信号とする例を用いると、蓄積エリア選択特徴データベース38は、各CM提供者名に対応付けて、前記予め計算された蓄積エリア選択特徴を記憶している。

なお、この例では、蓄積エリア選択特徴データベース38に、予め計算した蓄積エリア選択特徴を計算して蓄積しているが、蓄積映像信号(生の映像信号)を記憶するデータベースを設けるようにしても良い。

[0133] ここで、目的特徴エリア選択部35及び蓄積特徴エリア選択部36での閾値の設定について説明する。

上記閾値は、例えば、正規化特徴ベクトルの各要素の値と統計処理により求められた基準となる値(ここでは「0」)との差分の絶対値が大きい要素を選択する場合、正規化特徴ベクトルの各要素の値と「0」との差分の絶対値が最大となる値の例えば8割の数値を閾値とし、各要素の中で、差分の絶対値が閾値を越えた要素を選択することが考えられる。

[0134] 図6は、このことを示すものである。図6において、正規化特徴ベクトルの各要素d1～d6の値と「0」との差分の絶対値をD1～D6とする。ここで、差分の絶対値D1～D6



の中で最大となるものを選び、その値の例えば8割を閾値とする。

図6に示すように、差分の絶対値が最大となるのは、要素d2の差分の絶対値D2である。よって、D2との差分の絶対値の8割を閾値  $|\theta|$  として設定する。

$$|\theta| = 0.8 \times D2$$

そして、各要素d1～d6の値との差分の絶対値をD1～D6と、閾値  $|\theta|$  とをそれぞれ比較し、閾値を越えるものを抽出する。なお、選択されたものは丸印が付されている。

[0135] 図6では、要素d2の差分の絶対値D2と、要素d4の差分の絶対値D4と、要素d5の絶対値D5と、要素d6の差分の絶対値D6とが閾値  $|\theta|$  を越えているので、これらの要素が選択特徴のパターンとして選択される。

[0136] なお、ここでは、差分の絶対値が最大となるものの8割の値を閾値としているが、これは一例であり、8割に限定されるものではない。

[0137] また、要素毎の下限値を設けておき、この下限値(1つ又は複数個)を得る要素が選択されるまで、閾値を徐々に下げるように制御しても良い。例えば、図7の例では、上位2つの要素を選択する場合を示し、閾値を徐々に下げていくことにより、差分の絶対値が最大となる要素d2の値と「0」との差分の絶対値D2と、要素d5の値と「0」との差分の絶対値D5とが選択特徴のパターンとして選択される。

[0138] また、目的特徴エリア選択部35は、上記目的特徴における所定の範囲の特徴ベクトルの各要素の平均値を求め、各要素からこの平均値を除算した結果の絶対値が所定の閾値を超えた要素を、目的正規化特徴から選択し、選択された要素の多次元ベクトルからなる目的エリア選択特徴を導いても良い。

[0139] 同様に、蓄積特徴エリア選択部36は、上記蓄積特徴における所定の範囲の特徴ベクトルの各要素の平均値を求め、各要素からこの平均値を除算した結果の絶対値が所定の閾値を超えた要素を、蓄積正規化特徴から選択し、選択された要素の多次元ベクトルからなる蓄積エリア選択特徴を導いても良い。

[0140] 上記閾値は、例えば、正規化特徴ベクトルの各要素の値と平均値との差分の絶対値が大きい要素を選択する場合、正規化特徴ベクトルの各要素の値と平均値との差分の絶対値が最大となる値の例えば8割の数値を閾値とし、各要素の中で、平均値

との差分の絶対値が閾値を越えた要素を選択するようにしても良い。

このとき、要素の下限値を設けておき、この下限値を越える要素が選択されるまで、閾値を徐々に下げるように制御しても良い。

[0141] 更に、目的特徴エリア選択部35は、上記目的特徴及び目的正規化特徴を入力し、この目的特徴における所定の範囲(一定区画)の特徴ベクトルの要素毎の標準偏差を求め、次に、目的正規化特徴の対応する配列位置の特徴ベクトルの要素毎に、上記標準偏差を乗算して、乗算結果の絶対値を計算して統計量とし、この統計量において最上位又は最上位から複数個(例えば、2個)を、目的正規化特徴から選択して、特徴パターンである、選択された要素の多次元ベクトルからなる目的エリア選択特徴を導いても良い。

[0142] 同様に、蓄積特徴エリア選択部36は、上記蓄積特徴及び蓄積正規化特徴を入力し、この蓄積特徴における所定の範囲(一定区画)の特徴ベクトルの要素毎の標準偏差を求め、次に、蓄積正規化特徴の対応する配列位置の特徴ベクトルの要素毎に、上記標準偏差を乗算して、乗算結果の絶対値を計算して統計量とし、この統計量において最上位又は最上位から複数個(例えば、2個)を、目的正規化特徴から選択して、特徴パターンである、選択された要素の多次元ベクトルからなる目的エリア選択特徴を導いても良い。

[0143] 次に、図面を参照して本実施形態の映像信号探索システムの動作の説明を行う。

図8は、図5の映像信号探索システムの動作例を示すフローチャートである。

以下の説明では、1フレームを複数に分割した各領域のRGBの各色の平均値の情報をを用いた実施例を示す。

[0144] 図8において、蓄積特徴計算部32は、与えられた蓄積映像信号を読み込んで出力し(ステップU1)、入力した蓄積映像信号に対して特徴抽出を行う。

このとき、蓄積特徴計算部32は、蓄積映像の時系列データをもとにして、図9に示すように、例えば蓄積映像の1フレームの画像を横に3等分かつ縦に3等分して合計9の領域に分割し、それぞれの分割(領域)内におけるRGBの各色の画素について平均値を算出する。第i番目のフレームの場合、第1サブ画面のRGBに関する平均値を各々( $x_{i,1}$ ,  $x_{i,2}$ ,  $x_{i,3}$ )とし、第2サブ画面のRGBに関する平均値を各々( $x_{i,4}$ ,  $x_{i,5}$ ,  $x_{i,6}$ )とする。

WO 2006/006528

39

PCT/JP2005/012667

$x_{i,6}$ )とし、第3サブ画面のRGBに関する平均値を各々( $x_{i,7}$ ,  $x_{i,8}$ ,  $x_{i,9}$ )とし、以下同様にして、各サブ画面のRGBに関する平均値を各々要素とする多次元ベクトルを生成する。この多次元ベクトルを局所領域原色特徴と呼ぶ。

この場合、前述の $x_{i,1}$ 等は局所領域原色特徴の要素となる。

こうして得られた9個の領域におけるRGBそれぞれの平均画素値からなる合計27次元のベクトルを特徴ベクトルとして抽出し、蓄積特徴とする。この場合、特徴ベクトルは、1フレーム毎に得られる(ステップU2)。

また、蓄積特徴の抽出方法には、デジタル動画像圧縮技術であるMPEG (Moving Picture Coding Experts Group)符号化方式(動き補償とDCT (Discrete Cosine Transform)と可変長符号化からなる)やフーリエ変換の振幅成分を使用する方法がある。

[0145] 次に、蓄積特徴正規化部34は、蓄積特徴計算部32から蓄積特徴を読み込み、この蓄積特徴の特徴ベクトルの各要素毎に、所定の一定区画の平均値と標準偏差とを計算して求める。例えば、蓄積特徴正規化部34は、各領域における各RGB色毎にある時間区間の値から平均値と標準偏差とを求め、求められた平均値と標準偏差とを用いて正規化を行う。

このとき、蓄積特徴正規化部34による正規化後の局所領域原色特徴を正規化局所原色特徴と呼ぶ時、そのk番目の要素は、以下の(31)式となる。

[0146] [数31]

$$y_{i,k} = \frac{1}{\sigma_{i,k}} x_{i,k} - m_{i,k} \quad \dots (31)$$

[0147] 但し、(31)式における $x_{(i+j),k}$ は、全フレームに通し番号を付与した場合の(i+j)番目のフレームの局所領域特徴のk番目の要素を表す。jは設定された時間区間内のフレームにおける中心フレームに対する相対的な番号であり、設定された時間区間内に2M個のフレームが含まれる場合、 $-M \leq j \leq M-1$ 、jは整数となる。iは設定された時間区間のフレームのうち中心フレームの通し番号( $i \geq M$ )であり、j=0の時の

WO 2006/006528

40

PCT/JP2005/012667

通し番号に相当する。また、 $m_{i,k}$  は、 $-M \leq j \leq M-1$  の  $x_{(i+j),k}$  に対する平均値であり、(32)式で表される。 $\sigma_{i,k}$  は、 $-M \leq j \leq M-1$  の  $x_{(i+j),k}$  に対する標準偏差であり、以下に示す(33)式により求められる。

[0148] [数32]

$$m_{i,k} = \frac{1}{2M} \sum_{j=-M}^{M-1} x_{(i+j),k} \quad \cdots (32)$$

[0149] [数33]

$$\sigma_{i,k}^2 = \frac{1}{2M} \sum_{j=-M}^{M-1} (x_{(i+j),k} - m_{i,k})^2 \quad \cdots (33)$$

[0150] 局所時間内の統計処理を行うために、2Mフレームに相当する時間窓を設定したとすると、図10に示すように、2Mフレーム分の特徴ベクトルが得られる。特徴ベクトルは、図9に示したように、1フレームを9つの領域に分割し、それぞれの領域においてRGBの各画素について平均化したもので、27の要素からなる。よって、2Mフレームに相当する時間窓を設定すると、特徴ベクトルが2M個得られ、この2Mフレームからなる特徴ベクトルを纏めると、以下のような(2M)行N列の行列で表現できる。但し、Nは1フレーム毎に得られる特徴ベクトルの要素の数である。nを、1フレームを分割した際のサブ画面の数を表す整数とした時、Nは、 $N=3 \times n$ で表される。なお、前記時間窓のフレーム数は、2Mという偶数で記載したが、偶数に限られず、奇数であってもよい。

[0151] [数34]

[0155] ここで、蓄積特徴エリア選択部36は、得られた各要素と標準偏差との乗算結果である統計量 $z_{i,k}$  から、一定区画単位に要素毎に最上位から複数個の要素、例えば上位の2つの要素を選択する。統計量 $z_{i,k}$  を選択するための所定の区間(時間窓又はフレーム数)は、正規化局所領域原色特徴の要素 $y_{i,k}$  を算出した際の時間窓とは独立に決定でき、一致する必要はない。ここでは、前記時間窓に含まれるフレーム数を $M'$ と

WO 2006/006528

42

PCT/JP2005/012667

する。

つまり、局所時間内の特徴ベクトルは、式(36)に示すような行列で表現できる。この行列の各行の要素のうち、降順に上述した所定の個数に対応するP個(Pは正整数)、例えば、(P=2)個を選択する。

[0156] [数36]

$$Z = \begin{pmatrix} z_{i,1} & z_{i,2} & z_{i,3} & z_{i,4} & z_{i,5} & z_{i,6} & \cdots & z_{i,N} \\ z_{i+1,1} & z_{i+1,2} & z_{i+1,3} & z_{i+1,4} & z_{i+1,5} & z_{i+1,6} & \cdots & z_{i+1,N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ z_{i+(M'-1),1} & z_{i+(M'-1),2} & z_{i+(M'-1),3} & z_{i+(M'-1),4} & z_{i+(M'-1),5} & z_{i+(M'-1),6} & \cdots & z_{i+(M'-1),N} \end{pmatrix} \quad \cdots (36)$$

[0157] このような選択は、選択された( $M' \times P$ )個の要素に対応する要素が「0」であり、他の要素が( $-z_{i,k}$ )である行列(選択マスク)を加算することで実現できる。例えば、 $z_{i,1}$ ,  $z_{i,3}$ ,  $z_{i+1,2}$ ,  $z_{i+1,N}$ ,  $z_{i+(M'-1),2}$ ,  $z_{i+(M'-1),3}$ を選択する選択マスクは、以下のようになる。

[0158] [数37]

$$Z_0 = \begin{pmatrix} 0 & -z_{i,2} & 0 & -z_{i,4} & -z_{i,5} & -z_{i,6} & \cdots & z_{i,N} \\ -z_{i+1,1} & 0 & -z_{i+1,3} & -z_{i+1,4} & -z_{i+1,5} & -z_{i+1,6} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -z_{i+(M'-1),1} & 0 & 0 & -z_{i+(M'-1),4} & -z_{i+(M'-1),5} & -z_{i+(M'-1),6} & \cdots & -z_{i+(M'-1),N} \end{pmatrix} \quad \cdots (37)$$

[0159] (36)式に示す行列に、(37)式に示す選択マスクを加算すると、以下のように、蓄積エリア選択特徴の行列が得られる。

[0160] [数38]

$$Z_s = Z + Z_0 \quad \cdots (38)$$

$Z_s$ : 蓄積エリア選択特徴

WO 2006/006528

43

PCT/JP2005/012667

[0161] そして、蓄積特徴エリア選択部36は、上記一定区間において、蓄積正規化特徴から要素毎に選択された要素のベクトルからなる多次元ベクトルを蓄積エリア選択特徴として出力する(ステップU4)。

[0162] このとき、蓄積特徴エリア選択部36は、計算により得られた蓄積エリア選択特徴を、特徴照合部37へ直接送信するか、又は、一旦蓄積エリア選択特徴データベース38への登録のいずれかの処理を行う。

特徴照合部37がリアルタイムに蓄積エリア選択特徴と、目的エリア選択特徴との比較を行う場合、蓄積特徴エリア選択部36は、特徴照合部37へ、入力されている蓄積映像信号の蓄積エリア選択特徴を出力する。また、蓄積エリア選択特徴データベース38へ蓄積映像信号のデータを登録する場合、蓄積特徴エリア選択部36は特徴照合部37へ蓄積エリア選択特徴を送信せずに、蓄積エリア選択特徴データベース38へ、例えば映像広告提供者名、或いは番組名や映画の題名に対応させて蓄積エリア選択特徴の登録処理を行う。

[0163] 目的特徴計算部31は、与えられた目的映像信号を読み込んで入力し(ステップU5)、入力した目的映像信号に対して特徴抽出を行う。

このとき、目的特徴計算部31は、蓄積特徴計算部32と同様に、目的映像の時系列データをもとにして、例えば目的映像の1フレームの画像を横に3等分かつ縦に3等分して合計9の領域に分割し、それぞれの分割(領域)内におけるRGBの各色の画素について平均値を算出する。こうして得られた9個の領域におけるRGBそれぞれの平均画素値からなる合計27次元のベクトルを特徴ベクトルとして抽出し、目的特徴とする(ステップU6)。

また、蓄積特徴の抽出方法には、デジタル動画像圧縮技術であるMPEG符号化方式(動き補償とDCTと可変長符号化からなる)やフーリエ変換の振幅成分を使用する方法がある。

[0164] 次に、目的特徴正規化部33は、目的特徴計算部31から目的特徴を読み込み、この目的特徴の特徴ベクトルの要素毎に、所定の一定区画の平均値と標準偏差とを計算して求める。

即ち、目的特徴正規化部33は、蓄積特徴正規化部34と同様に、(31)～(33)式により得られた多次元ベクトルの時系列な配列( $k$ で示される順番)を目的正規化特徴とする(ステップU7)。

- [0165] 次に、目的特徴エリア選択部35は、目的特徴計算部31から目的特徴を、また、目的特徴正規化部33から、目的正規化特徴を読み込んで入力し、蓄積特徴エリア選択部36と同様に、目的特徴の一定区画から求めた標準偏差  $\sigma_{i,k}$  を、要素毎に乗算して乗算結果を統計量とする。同様のことは、目的特徴正規化部33から目的正規化特徴と標準偏差とを、目的特徴エリア選択部35に読み込み入力することでも達成できる。

そして、目的特徴エリア選択部35は、要素毎の配列において、上記統計量から数値の大きい順に上位2つの要素を選択して、目的正規化特徴から要素毎に選択された要素のベクトルからなる多次元ベクトルを目的エリア選択特徴として出力する(ステップU8)。

- [0166] 次に、特徴照合部37は、目的特徴エリア選択部35及び蓄積特徴エリア選択部36により得られた目的エリア選択特徴及び蓄積エリア選択特徴を各々読み込む。

また、特徴照合部37は、目的映像信号及び蓄積映像信号を同時に入力して、リアルタイムに類似を判定する場合以外、蓄積エリア選択特徴データベース38から順次比較を行う蓄積エリア選択特徴を読み出して、目的エリア選択特徴と比較する。

このとき、特徴照合部37は、蓄積エリア選択特徴において、目的特徴エリア選択部35で与えられた目的エリア選択特徴と同じ長さの特徴ベクトルの配列を照合区間として設定する。

即ち、特徴照合部37は、複数の上記一定区画からなる目的エリア選択特徴を照合区間とし、蓄積エリア選択特徴において、目的エリア選択特徴と同一の一定区画数からなる配列長を、目的エリア選択特徴に対する照合区間として対応させて、順次、両照合区間を類似度により比較する。

- [0167] そして、特徴照合部37は、目的エリア選択特徴と蓄積エリア選択特徴とにおける照合区間の比較箇所との類似度判断数を下記(39)式を用いて計算する。特徴照合部37で用いる照合区間(時間窓又はフレーム数)は、蓄積エリア選択部36又は目的エ



WO 2006/006528

45

PCT/JP2005/012667

リア選択部35で用いた時間区間(時間窓)及び蓄積又は目的特徴正規化部33で用いた時間区間(時間窓)とは独立に決定することができる、即ち、必ずしも一致した時間区間を使用する必要はない。

特徴照合部37は、目的エリア選択特徴及び蓄積エリア選択特徴の照合区間における各々の正規化局所領域原色特徴 $y_{q\_i,k}$ と、 $y_{s\_i,k}$ との一致箇所のユークリッド距離と、非一致箇所の数に所定の定数 $a$ を乗じた値の和を計算し、これを類似度判断数 $S$ として出力する(ステップU9)。

なお、蓄積エリア選択特徴と目的エリア選択特徴とを生成する際の各々の選択マスクにおいて、同じ $i, k$ を有する要素(同じ位置の要素)が「0」である要素を「一致」といい、それ以外を「非一致」という。

[0168] [数39]

$$S = \sum_{k=q \cap s} (y_{q\_i,k} - y_{s\_i+r,k})^2 + a (\max(|q|, |s|) - |q \cap s|) \quad \dots (39)$$

$x \cap y$ :  $x$ と $y$ の一致する要素

[0169] 但し、 $q, s$ は各々目的特徴のエリア選択箇所、蓄積特徴のエリア選択箇所であることを意味し、 $r$ は整数であり、目的特徴と蓄積特徴とにおけるフレーム番号が等しくなくとも良いことを表す変数である。 $a$ は例えば $y_{i,k}$ のユークリッド距離の差分値のとり得る値の最大値、即ち、 $y_{i,k}$ の量子化レベル数の2乗とすることが考えられる。 $x$ はエリア選択箇所 $x$ の要素数、 $\max(x, y)$ は、次式で定義する。

[0170] [数40]

$$\max(x, y) = \begin{cases} x & (x \geq y) \\ y & (\text{otherwise}) \end{cases} \quad \dots (310)$$

[0171] ここで、例えば、特徴照合部37は目的エリア選択特徴が15秒の長さであるとする、配列の要素としての特徴ベクトルを0.1秒間隔に合計150箇所抽出する。各サンプリングにおいてフレーム画像を分割し、RGB号それぞれの平均画素値を求めてい

WO 2006/006528

46

PCT/JP2005/012667

るため、これらのベクトルからなる $(150 \times 27)$ の4050次元から、各領域における各RGB色強度の上位2つ(目的特徴エリア選択部35が抽出)を用い、54の要素からなる多次元ベクトルを照合に用いる目的ベクトルとする。

[0172] また、特徴照合部37は、上記目的ベクトルと同様に、蓄積エリア選択特徴のデータの先頭から、15秒の長さを単位に順次照合区間として設定して、特徴ベクトルの配列から、特徴ベクトルを0.1秒間隔に合計150箇所抽出し、各サンプリングにおいてフレーム画像を分割し、RGBそれぞれの平均画素値を求めているため、それらからなる $(150 \times 27)$ の4050次元から、各領域における各RGB色強度の上位2つ(蓄積特徴エリア選択部36が抽出)を用い、54の要素からなる多次元ベクトルを照合に用いる蓄積ベクトルとする。

[0173] そして、特徴照合部37は、目的エリア選択特徴と蓄積エリア選択特徴との照合を行う。

即ち、(39)式を用いて、上記目的ベクトルと蓄積ベクトルとの類似度判断数Sを計算し、予め設定した探索閾値と比較処理を行う(ステップU10)。照合の結果、(39)式で求められる類似度判断数Sが予め設定された探索閾値より大きい(類似度が小さい)場合には、時間窓をシフトして、フィードバックする(ステップU11)。

[0174] ステップU11で、照合区間を蓄積エリア選択特徴の先頭から順次ずらしながら(例えば、1フレームずつシフトさせながら)、上記目的ベクトルと蓄積ベクトルとの類似度判断数Sを計算し、予め設定した探索閾値と比較処理を行い、目的エリア選択特徴と蓄積エリア選択特徴とを、蓄積エリア選択特徴の最後まで照合処理した後、照合区間毎の類似度判断数Sが予め設定された探索閾値より低い蓄積エリア選択特徴の照合区間の領域が検索されたら、この照合区間を探索結果として出力する(ステップU12)。

なお、照合する際の目的映像信号のフレーム数と蓄積映像信号のフレーム数とは必ずしも等しい必要はない。

[0175] 以上説明したように、本発明の実施形態では、蓄積映像信号及び目的映像信号から統計的に特徴的な要素のみを選択して照合を行っている。

[0176] なお、特徴照合部37は、探索結果として、比較した結果において、もっとも類似度

WO 2006/006528

47

PCT/JP2005/012667

判断数が低い蓄積エリア選択特徴の照合区間の領域を出力するようにしても良い。

- [0177] また、特徴照合部37は、探索閾値を複数の照合区間の類似度が下回った場合、類似度の上位(低い数値のものから)N番目までの照合区間を出力するようにすることも可能である。

更に、特徴照合部37は、探索閾値を下回る照合区間がない場合など、該当箇所がないことを示す情報を通知し、新たな蓄積映像信号の蓄積エリア選択特徴を、蓄積エリア選択特徴データベース38から読み出し、上記探索閾値以下の照合区間を有する蓄積エリア選択特徴が探索されるまで、ステップU9以降の探索処理を継続させて行うようにしても良い。

- [0178] また、図5における目的特徴計算部31、目的特徴正規化部33及び目的特徴エリア選択部35を各ユーザの端末(例えばパーソナルコンピュータ)にインストールしておき、映像配信を行うサービスプロバイダに蓄積特徴計算部32、蓄積特徴正規化部34、蓄積特徴エリア選択部36、特徴照合部37及び蓄積エリア選択特徴データベース38を有する映像信号探索サーバを設けても良い。

これにより、ユーザがビデオカメラ付携帯電話等で受信した映像信号を目的映像信号として、目的エリア選択特徴まで生成し、この目的エリア選択特徴を上記映像信号探索サーバへ、インターネットなどを介して送信し、この目的エリア選択特徴に類似した蓄積映像信号を探索してもらうように要求する構成とする。このとき、蓄積エリア選択特徴及び目的エリア選択特徴を計算するときの一定区画の特徴ベクトルの配列長等の規定は、映像信号探索サーバと端末とにおいて予め一致させておく。

- [0179] 次に、上記映像信号探索システムを適用した実施形態を示す。本発明の映像信号探索システムは、実環境で収録した反射や途切れのある断片的な映像信号を使って、一致する映像を探索し情報検索を行うのに利用できる。例えば、街頭の大型スクリーンに映し出されるCMをユーザがビデオカメラ付携帯電話で撮影し、映像信号探索サービスなどに送信する。映像信号探索サービス提供者は、この映像信号に一致又は類似する映像をデータベースより検索し、コンテンツ或いはこの映像に関する情報(例えば商品情報、製品特徴、サービス、撮影地、出演者、ホームページ等)を有料、又は無料でユーザに提供するといった構成が可能である。

WO 2006/006528

48

PCT/JP2005/012667

ここで、検索する映像の入力方法においては、ビデオカメラ付端末のファインダ又はスクリーンを撮りたい映像(目的映像信号)のフレームに合わせて撮影するか、又は撮影された映像中の動画フレームをメニュー操作、或いはペン入力等によるマニュアルレースで範囲指定することが望ましい。

更に、本発明により従来は正しく探索することの難しかった、家庭用ビデオでダビングを反復した映像や、ビットレートの低い映像も容易に探索できるため、インターネット上の動画の著作権管理システムや、CM情報検索サービスなどにも広く応用可能である。

他には、ユーザがビデオカメラで撮影した動画イメージを使って、そのイメージに近いクリップを、配信されている映画やテレビプログラムから切り取って編集したりすることにも利用できるであろう。

[0180] <第4の実施の形態>

以下、本発明の第4の実施の形態について図面を参照しながら説明する。図11は、本発明の一実施形態であり、映像信号を対象とする反射や途切れに頑健な映像信号探索システムの構成を示すブロック図である。

この図11に示す映像信号探索システムは、映像信号を対象とする反射や途切れに頑健な信号検出を実現するものであり、目的特徴計算部41と、蓄積特徴計算部42と、目的特徴正規化部43と、蓄積特徴正規化部44と、目的特徴量子化部45と、蓄積特徴量子化部46と、特徴照合部47と、蓄積量子化特徴データベース48とで構成されており、蓄積映像信号、即ち検索される映像信号と、目的映像信号、即ち検索したい映像信号を入力し、目的時系列信号に類似した蓄積時系列信号中の箇所を出力する。

[0181] 目的特徴計算部41は、目的映像信号を入力とし、この目的映像信号を所定の間隔で時系列にサンプリングして得られる離散データである蓄積時系列信号から、抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる目的特徴を算出するものである。

[0182] 蓄積特徴計算部42は、目的特徴計算部41と同様に、蓄積映像信号を所定の間隔で時系列にサンプリングして得られる離散値である蓄積時系列信号から、例えば、蓄

WO 2006/006528

49

PCT/JP2005/012667

積映像信号からサンプリング単位に周波数毎のパワースペクトル値等を抽出して特徴量とし、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる蓄積特徴を算出するものである。

[0183] 目的特徴正規化部43は、上記目的特徴から、その目的特徴に隣接した目的特徴を含む周辺の複数の目的特徴から導いた統計量(目的統計量とも呼ぶ)を用いて、特徴ベクトルの要素毎に独立に正規化し、正規化された値からなる特徴ベクトルを有する目的正規化特徴を導くものである。

[0184] 蓄積特徴正規化部44は、上記蓄積特徴から、その蓄積特徴に隣接した蓄積特徴を含む周辺の複数の蓄積特徴から導いた統計量(蓄積統計量とも呼ぶ)を用いて特徴ベクトルの要素毎に独立に正規化し、正規化された値からなる特徴ベクトルを有する蓄積正規化特徴を導くものである。

[0185] ここで、「蓄積特徴に『隣接した』蓄積特徴」の表現で用いられる「隣接した」とは、例えば、映像信号を時刻に関してサンプリングした際に得られる離散時刻のフレームに対し、その前後の離散時刻のフレームを表現する場合の「前後の離散時刻の」に対応する。

また、「周辺の複数の」とは、前記例を用いると、「ある離散時刻の間の」を意味する。

[0186] 目的特徴量子化部45は、上記目的正規化特徴の各要素において、同一若しくは個別の閾値によりスカラー量子化し、この量子化された値をベクトルの要素とする多次元ベクトルとして目的量子化特徴を導くものである。

[0187] 蓄積特徴量子化部46は、上記蓄積正規化特徴の各要素において、同一若しくは個別の閾値によりスカラー量子化し、この量子化された値をベクトルの要素とする多次元ベクトルとした蓄積量子化特徴を導くものである。

なお、映像信号における蓄積特徴についての縮退させたベクトルおよび、目的特徴についての縮退させたベクトルとは、上述と同様の計算により得られるものである。

[0188] 量子化を行う上記所定の閾値は、目的及び蓄積正規化特徴を2値化する場合、各符号の代表数値ともとの要素の数値との平均2乗誤差が最小化される値を使用する。代表数値としては、例えば、2つの符号のうち一方を閾値よりも1大きい値を使用し、もう一方を閾値よりも1小さい値を使用する。

WO 2006/006528

50

PCT/JP2005/012667

[0189] 特徴照合部47は、上記蓄積量子化特徴に所定の範囲を照合区間として設定し、この照合区間と同一の長さの被照合区間を目的量子化特徴に設定して、上記照合区間と被照合区間との類似度を演算し、予め設定された探索閾値と比較して、類似しているか否かの判定を行う。

また、特徴照合部47は、上記照合区間の照合処理が終了すると、新たな照合区間を設定するため、配列の隣接する同一時間幅の範囲に照合区間をずらす処理を行っている。

[0190] 蓄積量子化特徴データベース48は、蓄積特徴計算部42、蓄積特徴正規化部44及び蓄積特徴量子化部46により、予め蓄積量子化特徴を計算し、記憶している。複数のCM(Commercial Message)映像を蓄積映像信号とする例を用いると、蓄積量子化特徴データベース48は、各CM提供者名に対応付けて、予め計算された蓄積量子化特徴を記憶している。

なお、この例では、蓄積量子化特徴データベース48に、予め計算した蓄積量子化特徴を計算して蓄積しているが、蓄積映像信号(生の映像信号)を記憶するデータベースを設けるようにしても良い。

[0191] 次に、図面を参照して本実施形態の映像信号探索システムの動作の説明を行う。図12は、図11の映像信号探索システムの動作例を示すフローチャートである。

なお、以下の説明では、1フレームを複数に分割した各領域のR(赤)G(緑)B(青)の各色の平均値の情報を用いた実施例を示す。

[0192] 図12において、蓄積特徴計算部42は、与えられた蓄積映像信号を読み込んで出力し(ステップV1)、入力した蓄積映像信号に対して特徴抽出を行う。

このとき、蓄積特徴計算部42は、蓄積映像の時系列データをもとにして、蓄積映像の1フレームの画像をn領域(例えば横に3等分、縦に3等分した合計9の領域)に分割し、それぞれの分割領域内におけるRGBの各色の画素について平均値を算出する。

[0193] つまり、図13に示すように、フレーム(-M)からフレーム(M-1)の2Mフレームに相当する時間窓を設定したとすると、その間にあるフレーム0(0は時間窓内のフレームにおける相対番号を表し、0番目のフレームの意味である。)において、図14に示

WO 2006/006528

51

PCT/JP2005/012667

すように、1フレームをn個のサブ画面に分割する。フレーム0の第1サブ画面のRGBの各色の画素を平均して平均値( $x_{i,1}, x_{i,2}, x_{i,3}$ )を取得し、第2サブ画面のRGBの各色の画素を平均して平均値( $x_{i,4}, x_{i,5}, x_{i,6}$ )を取得し、第3サブ画面のRGBの各色の画素を平均して平均値( $x_{i,7}, x_{i,8}, x_{i,9}$ )を取得し、以下同様にして、第nサブ画面のRGBの各色の画素を平均して平均値( $x_{i,(3n-2)}, x_{i,(3n-1)}, x_{i,3n}$ )を取得する。但し、第1番目の添え字は全フレームに付与された通し番号を示し、第2番目の添え字は、そのフレーム内での要素の番号を示す。また、第1番目の添え字は、全フレームに付与された通し番号iと、第i番目のフレームに対する相対的な位置を表すjとの和( $i+j$ )で表すこともある。

- [0194] これにより、例えば、 $n=9$ とすると、各フレームにおいて( $3 \times 9=27$ )次元の値が得られ、こうして得られた27次元の多次元ベクトルを特徴ベクトルとして抽出して蓄積特徴とする。(ステップV2)。前記多次元ベクトルを局所領域原色特徴と呼ぶ。この場合、前述の $x_{i,1}$ 等は局所領域原色特徴の要素となる。

また、蓄積特徴の抽出方法には、デジタル動画像圧縮技術であるMPEG (Moving Picture Coding Experts Group) 符号化方式(動き補償とDCT(Discrete Cosine Transform)と可変長符号化からなる)やフーリエ変換の振幅成分を使用する方法がある。

- [0195] 次に、蓄積特徴正規化部44は、蓄積特徴計算部42から蓄積特徴を読み込み、この蓄積特徴の特徴ベクトルの要素ごとに、所定の一定区画の平均値と標準偏差とを計算して求める。

例えば、蓄積特徴正規化部44は、各領域における各RGB色毎にある時間区間の値から平均値と標準偏差とを求め、求められた平均値と標準偏差とを用いて正規化を行う。

このとき、蓄積特徴正規化部44による正規化後の局所領域原色特徴を正規化局所原色特徴と呼ぶ時、そのk番目の要素 $y_{i,k}$ は、以下の(41)式となる。

- [0196] [数41]

$$y_{i,k} = \frac{1}{\sigma_{i,k}} (x_{i,k} - m_{i,k}) \quad \dots (41)$$

WO 2006/006528

52

PCT/JP2005/012667

[0197] [数42]

$$m_{i,k} = \frac{1}{2M} \sum_{j=-M}^{M-1} x_{(i+j),k} \quad \cdots (42)$$

[0198] [数43]

$$\sigma^2_{i,k} = \frac{1}{2M} \sum_{j=-M}^{M-1} (x_{(i+j),k} - m_{i,k})^2 \quad \cdots (43)$$

[0199] 但し、(42)式における $x_{(i+j),k}$ は、全フレームに通し番号を付与した場合の $(i+j)$ 番目のフレームの局所領域特徴の $k$ 番目の要素を表す。 $j$ は設定された時間区間内のフレームにおける中心フレームに対する相対的な番号であり、設定された時間区間内に $2M$ 個のフレームが含まれる場合、 $-M \leq j \leq M-1$ 、 $j$ は整数となる。 $i$ は設定された時間区間のフレームのうちの中心フレームの通し番号( $i \geq M$ )であり、 $j=0$ の時の通し番号に相当する。また、 $m_{i,k}$ は、 $-M \leq j \leq M-1$ の $x_{(i+j),k}$ に対する平均値である。(43)式における $\sigma_{i,k}$ は $-M \leq j \leq M-1$ の $x_{(i+j),k}$ に対する標準偏差である。

[0200] 或るフレームを中心フレームとし、その中心フレームの前後のフレームの局所領域原色特徴を用いて統計処理(正規化)を行い、その中心フレームに対して正規化局所領域原色特徴を得る。1フレームずつずらしながら中心フレームを設定することで、全てのフレームに対して正規化局所領域原色特徴を得ることができる。この工程は、局所領域原色特徴を正規化局所領域特徴へ写像する工程である。蓄積特徴の前記写像を正規化蓄積特徴 $Y$ と呼ぶ(ステップV3)。

[0201] [数44]

$$Y = (y_{i,1}, y_{i,2}, y_{i,3}, \cdots, y_{i,N}) \quad \cdots (44)$$



WO 2006/006528

53

PCT/JP2005/012667

[0202] 但し、Nは1フレーム毎に得られる特徴ベクトルの要素の数である。1フレームをn分割した際、Nは、 $N=3 \times n$ で表される。

次に、蓄積特徴量子化部46は、蓄積特徴正規化部44から、蓄積特徴を読み込んで入力し、蓄積正規化特徴の多次元ベクトルの要素毎に、1以上の閾値により、量子化する。

[0203] 例えば、閾値を $t_1, t_2, \dots, t_R$ とし、上記(44)式で示されるベクトルYを量子化すると、以下のように、R個の閾値を用いた場合には、 $y_{i,k}$ は、(R+1)元符号化される。

[0204] [数45]

$$z_{i,k} = \begin{cases} 0 & y_{i,k} \leq t_1 \\ 1 & t_1 < y_{i,k} \leq t_2 \\ \vdots & \vdots \\ R-1 & t_{R-1} < y_{i,k} \leq t_R \\ R & t_R < y_{i,k} \end{cases} \quad \dots (45)$$

[0205] 以降、単純化のために、閾値tにより2元符号化(2値化)される場合について説明する。閾値tによる2元符号化は、以下のようになる。

[0206] [数46]

$$z_{i,k} = \begin{cases} 0 & y_{i,k} \leq t \\ 1 & y_{i,k} > t \end{cases} \quad \dots (46)$$

[0207] 式(44)で示される行列を、閾値tにより2元符号化したものを、量子化特徴Zとすると、量子化特徴Zは、以下のように示される。

[0208] [数47]

WO 2006/006528

54

PCT/JP2005/012667

$$Z = (z_{i,1}, z_{i,2}, z_{i,3}, \dots, z_{i,N}) \quad \dots (47)$$

[0209] 蓄積特徴量子化部46は、この多次元ベクトルZ、又は、多次元ベクトルの時系列（即ち、複数のフレーム分のZ）を1つの新たな多次元ベクトルとしたもの、を蓄積量子化特徴として出力する（ステップV4）。

このとき、蓄積特徴量子化部46は、演算により得られた蓄積量子化特徴を、特徴照合部47へ直接送信するか、又は、一旦蓄積量子化特徴データベース48への登録のいずれかの処理を行う。

[0210] 特徴照合部47がリアルタイムに蓄積量子化特徴と、目的量子化特徴との比較を行う場合、蓄積特徴量子化部46は、特徴照合部47へ、入力されている蓄積映像信号の蓄積量子化特徴を出力する。また、蓄積量子化特徴データベース48へ蓄積映像信号のデータを登録する場合、蓄積特徴量子化部46は特徴照合部47へ蓄積量子化特徴を送信せずに、蓄積量子化特徴データベース48へ、例えば、映像広告提供者名、或いは番組名や映画の題名に対応させて蓄積量子化特徴の登録処理を行う。

[0211] 目的特徴計算部41は、与えられた目的映像信号を読み込んで入力し（ステップV5）、入力した目的映像信号に対して特徴抽出を行う。

このとき、目的特徴計算部41は、蓄積特徴計算部42と同様に、目的映像の時系列データをもとにして、例えば目的映像の1フレームの画像をn領域（例えば横に3等分、縦に3等分した合計9の領域）に分割し、それぞれの分割領域内におけるRGBの各色の画素について平均値を算出する。こうして得られた各領域におけるRGBそれぞれの平均画素値からなる例えば27次元のベクトル（n=9の場合）を特徴ベクトルとして抽出し、目的特徴とする（ステップV6）。

また、目的特徴の抽出方法には、デジタル動画像圧縮技術であるMPEG符号化方式（動き補償とDCTと可変長符号化からなる）やフーリエ変換の振幅成分を使用する方法がある。

WO 2006/006528

55

PCT/JP2005/012667

[0212] 次に、目的特徴正規化部43は、目的特徴計算部41から目的特徴を読み込み、この目的特徴の特徴ベクトルの要素ごとに、所定の一定区画の平均値と標準偏差とを演算して求める。

即ち、目的特徴正規化部43は、蓄積特徴正規化部44と同様に、(41)～(43)式により得られた多次元ベクトルの時系列な配列(kで示される順番)を目的正規化特徴とする(ステップV7)。

[0213] 次に、目的特徴量子化部45は、目的特徴正規化部43から、目的正規化特徴を読み込んで入力し、(44)式に示すような行列を生成し、この行列の要素ごとに、予め与えられた閾値を境に、蓄積特徴量子化部46と同様に、(46)式を用いて例えば二値に量子化して、(47)式に示されるように量子化特徴を得る。そして、この量子化特徴を目的量子化特徴として出力する(ステップV8)。

[0214] 次に、特徴照合部47は、目的特徴量子化部45及び蓄積特徴量子化部46により得られた目的量子化特徴及び蓄積量子化特徴を各々読み込む。

また、特徴照合部47は、目的映像信号及び蓄積映像信号を同時に入力して、リアルタイムに類似を判定する場合以外、蓄積量子化特徴データベース48から順次比較を行う蓄積量子化特徴を読み出して、目的量子化特徴と比較する。

このとき、特徴照合部47は、蓄積量子化特徴において、目的特徴量子化部45で与えられた目的量子化特徴と同じ長さの特徴ベクトルの配列を照合区間として設定する。

[0215] そして、特徴照合部47は、目的量子化特徴と蓄積量子化特徴との上記照合区間での類似度を演算する。類似度として、双方の特徴ベクトル間のハミング距離を演算する(ステップV9)。

特徴照合部47は、目的量子化特徴の特徴ベクトルの配列から、複数箇所の要素を抽出し、配列の要素としての特徴ベクトルとする。

例えば、目的量子化特徴が15秒の長さであるとする、配列の要素としての特徴ベクトルを0.1秒間隔に合計150箇所抽出する。各サンプリングにおいてフレーム画像を分割し、RGBそれぞれの平均画素値を求めているため、これらのベクトルからなる(150×27の)4050次元のベクトルを特徴照合部47で照合に用いる目的ベクトルと

WO 2006/006528

56

PCT/JP2005/012667

する。

[0216] また、特徴照合部47は、上記目的ベクトルと同様に、蓄積量子化特徴のデータの先頭から、15秒の長さを単位に順次照合区間として設定して、特徴ベクトルの配列から、特徴ベクトルを0.1秒間隔に合計150箇所抽出し、各サンプリングにおいてフレーム画像を分割し、RGBそれぞれの平均画素値を求めているため、このベクトルからなる(150×27の)4050次元のベクトルを照合に用いる蓄積ベクトルとする。

このとき、特徴照合部47が目的量子化特徴の特徴ベクトルの配列から、複数箇所の要素を抽出して目的ベクトルとするのであれば、目的特徴量子化部45において、予め、配列の要素としての特徴ベクトルを目的正規化特徴から抽出して、即ち0.1秒間隔に合計150箇所抽出して、スカラー量子化を行い、目的ベクトルとして特徴照合部47へ出力するようにしても良い。

[0217] 特徴照合部47は、照合区間を蓄積量子化特徴の先頭から順次ずらしながら、目的量子化特徴との、即ち、上記目的ベクトルと蓄積ベクトルとの、ハミング距離を演算することにより比較処理を行う(ステップV10)。

そして、特徴照合部47は、蓄積量子化特徴を最後まで照合した後、照合区間毎のハミング距離を参照して、ハミング距離が最も小さい照合区間の領域を探索結果として出力する(ステップV11)。

なお、特徴照合部47は、予めハミング距離の探索閾値が与えられていた場合、探索結果として、この探索閾値と選択された照合区間のハミング距離とを判定して(ステップV10)、この探索閾値を下回るもののみを出力することもできる(ステップV11)。

[0218] 但し、(R+1)元符号化した場合のように、非二次元の場合は、次に定義するリー距離を、ハミング距離の代わりに適用しても良い。但し、次式(48)において、 $\|z_{q_i, k} - z_{s_r, k}\|$  は、 $(z_{q_i, k} - z_{s_r, k}) \bmod (R+1)$  と、 $(z_{s_r, k} - z_{q_i, k}) \bmod (R+1)$  のうちいずれか小さい方を意味する。rはフレーム番号が、qとsとで必ずしも一致する必要はないことを表すために導入した記号である。

[0219] [数48]

$$d_L = \sum_{(i, r)} \sum_{k=1}^N \|z_{q_i, k} - z_{s_r, k}\| \quad \dots (48)$$

WO 2006/006528

57

PCT/JP2005/012667

[0220] 以上説明したように、本発明の実施形態では、蓄積映像信号及び目的映像信号から多次元ベクトルからなる蓄積特徴及び目的特徴として抽出し、この多次元ベクトルを量子化して量子化特徴を算出し、この蓄積量子化特徴と目的量子化特徴とをハミング距離等を用いて照合するようにしている。

[0221] また、特徴照合部47は、探索閾値を複数の照合区間のハミング距離が下回った場合、ハミング距離の上位(低い数値のものから)P番目(Pは所定の正整数)までの照合区間を出力するようにすることも可能である。

更に、特徴照合部47は、探索閾値を下回る照合区間がない場合など、該当箇所がないことを示す情報を通知し、新たな蓄積映像信号の蓄積量子化特徴を、蓄積量子化特徴データベース48から読み出し、上記探索閾値以下の照合区間を有する蓄積量子化特徴が探索されるまで、ステップV9以降の探索処理を継続させて行うようにしても良い。

[0222] また、図11における目的特徴計算部41、目的特徴正規化部43及び目的特徴量子化部45を各ユーザの端末(例えばパーソナルコンピュータ)にインストールしておき、映像配信を行うサービスプロバイダに蓄積特徴計算部42、蓄積特徴正規化部44、蓄積特徴量子化部46、特徴照合部47及び蓄積量子化特徴データベース48を有する映像信号探索サーバを設けても良い。

これにより、ユーザがビデオカメラ付携帯電話等で受信した映像信号を目的映像信号として、目的量子化特徴まで生成し、この目的量子化特徴を上記映像信号探索サーバへ、インターネットなどを介して送信し、この目的量子化特徴に類似した蓄積映像信号を探索してもらうように要求する構成とする。このとき、蓄積量子化特徴及び目的量子化特徴を計算するときの一定区画の特徴ベクトルの配列長等の規定は、映像信号探索サーバと端末とにおいて予め一致させておく。

[0223] 次に、上記映像信号探索システムを適用した実施形態を示す。本発明の映像信号探索システムは、実環境で収録した反射や途切れのある断片的な映像信号を使って、一致する映像を探索し情報検索を行うのに利用できる。例えば、街頭の大型スクリーンに映し出されるCMをユーザがビデオカメラ付携帯電話で撮影し、映像信号探

WO 2006/006528

58

PCT/JP2005/012667

索サービスなどに送信する。映像信号探索サービス提供者は、この映像信号に一致又は類似する映像をデータベースより検索し、コンテンツ或いはこの映像に関する情報(例えば商品情報、製品特徴、サービス、撮影地、出演者、ホームページ等)を有料、又は無料でユーザに提供するといった構成が可能である。

ここで、検索する映像の入力方法においては、ビデオカメラ付端末のファインダ又はスクリーンを撮りたい映像(目的映像信号)のフレームに合わせて撮影するか、又は撮影された映像中の動画フレームをメニュー操作、或いはペン入力等によるマニュアルレースで範囲指定することが望ましい。

更に、本発明により従来は正しく探索することの難しかった、家庭用ビデオでダビングを反復した映像や、ビットレートの低い映像も容易に探索できるため、インターネット上の動画の著作権管理システムや、CM情報検索サービスなどにも広く応用可能である。

他には、ユーザがビデオカメラで撮影した動画イメージを使って、そのイメージに近いクリップを、配信されている映画やテレビプログラムから切り取って編集したりすることにも利用できるであろう。

[0224] <第5の実施の形態>

図15は、本発明の第5の実施の形態を表しており、雑音やひずみに頑健な高精度信号検出システムの構成を示すブロック図である。ここで、信号とは時系列データであり、所定の再生装置があれば再生することが可能な情報全般を意味している。例えば音楽で言えば、音楽そのものであるアナログデータ、CD(Compact Disc)に記録されているデータ、WAVファイルのデータ、MP3(MPEG-1 Audio Layer3)ファイルのデータ等を示す。

また、映像で言えば、映像そのものであるアナログデータ、MPEG-2(Moving Picture Coding Experts Group 2)に代表されるデジタルデータ等を示す。ここでは、信号検出の対象を映像信号及び音響信号として説明を行う。

[0225] この図15に示す信号検出システムは、時系列信号を対象とした雑音やひずみに頑健で高精度な信号検出を実現するものであり、目的特徴計算部51と、蓄積特徴計算部52と、目的特徴正規化部53と、蓄積特徴正規化部54と、目的特徴エリア選択部5

WO 2006/006528

59

PCT/JP2005/012667

5と、蓄積特徴エリア選択部56と、目的特徴非線形量子化部57と、蓄積特徴非線形量子化部58と、特徴照合部59と、蓄積非線形量子化特徴データベース510とで構成され、蓄積時系列信号(蓄積信号)即ち検索したい信号と、目的時系列信号(目的信号)即ち検索される信号を入力し、目的時系列信号との類似した蓄積時系列信号中の箇所を出力する。なお、目的信号に基づいて、蓄積信号中から検索を行うことから、蓄積信号より目的信号の方が短くなる。

[0226] 目的特徴計算部51は、目的信号を所定の間隔で時系列にサンプリングして得られる離散値である目的時系列信号から、例えば、音響信号の場合、サンプリング単位に周波数毎のパワースペクトル値を抽出し、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる目的特徴を導くものである。また、映像信号の場合、フレームを分割し、分割領域内の画素の平均値を抽出し、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる目的特徴を導くものである。

[0227] 蓄積特徴計算部52も、目的特徴計算部51と同様に、蓄積信号を所定の間隔で時系列にサンプリングして得られる離散値である蓄積時系列信号から、例えば、音響信号の場合、サンプリング単位に周波数毎のパワースペクトル値を抽出し、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる蓄積特徴を導くものである。また、映像信号の場合、フレームを分割し、分割領域内の画素の平均値を抽出し、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる蓄積特徴を導くものである。

[0228] 目的特徴正規化部53は、上記目的特徴から、その目的特徴に隣接した目的特徴を含む周辺の複数の目的特徴から導いた統計量を用いて特徴ベクトルの要素毎に独立に正規化し、正規化された値からなる多次元ベクトルを有する目的正規化特徴を導くものである。

[0229] 蓄積特徴正規化部54は、上記蓄積特徴から、その蓄積特徴に隣接した蓄積特徴を含む周辺の複数の蓄積特徴から導いた統計量を用いて特徴ベクトルの要素毎に。独立に正規化し、正規化された値からなる多次元ベクトルを有する蓄積正規化特徴を導くものである。

WO 2006/006528

60

PCT/JP2005/012667

- [0230] 目的特徴エリア選択部55は、上記目的正規化特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、目的正規化特徴から選択し、選択された要素の多次元ベクトルからなる目的エリア選択特徴を導くものである。
- [0231] 蓄積特徴エリア選択部56は、上記蓄積正規化特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、蓄積正規化特徴から選択し、選択された要素の多次元ベクトルからなる蓄積エリア選択特徴を導くものである。
- [0232] ここで、図16及び図17を参照しつつ目的特徴エリア選択部55及び蓄積特徴エリア選択部56での閾値の設定について説明する。
- 例えば、正規化特徴ベクトルの各要素の値と統計処理により求められた基準となる値(ここでは「0」と)との差分の絶対値が大きい要素を選択する場合、正規化特徴ベクトルの各要素の値と「0」との差分の絶対値が最大となる値の例えば8割の数値を閾値とし、各要素の中で、差分の絶対値が閾値を越えた要素を選択することが考えられる。
- [0233] 図16において、正規化特徴ベクトルの各要素d1～d6の値と「0」との差分の絶対値をD1～D6とする。ここで、差分の絶対値D1～D6の中で最大となるものを選び、その値の例えば8割を閾値とする。
- 図16に示すように、差分の絶対値が最大となるのは、要素d2の差分の絶対値D2である。よって、D2との差分の絶対値の8割を閾値  $|\theta|$  として設定する。
- $$|\theta| = 0.8 \times D2$$
- そして、各要素d1～d6の値との差分の絶対値をD1～D6と、閾値  $|\theta|$  とをそれぞれ比較し、閾値を越えるものを抽出する。なお、選択されたものは丸印が付されている。
- [0234] 図16では、要素d2の差分の絶対値D2と、要素d4の差分の絶対値D4と、要素d5の絶対値D5と、要素d6の差分の絶対値D6とが閾値  $|\theta|$  を越えているので、これらの要素が選択特徴のパターンとして選択される。
- [0235] なお、ここでは、差分の絶対値が最大となるものの8割の値を閾値としているが、これは一例であり、8割に限定されるものではない。
- [0236] また、要素毎の下限値を設けておき、この下限値(1つ又は複数個)を得る要素が



WO 2006/006528

61

PCT/JP2005/012667

選択されるまで、閾値を徐々に下げるように制御しても良い。例えば、図17の例では、上位2つの要素を選択する場合を示し、閾値を徐々に下げていくことにより、差分の絶対値が最大となる要素d2の値と「0」との差分の絶対値D2と、要素d5の値と「0」との差分の絶対値D5とが選択特徴のパターンとして選択される。

[0237] また、目的特徴エリア選択部55は、上記目的特徴において、配列における所定の範囲の特徴ベクトルの各要素の平均値を求め、各要素からこの平均値を除算した結果の絶対値が所定の閾値を超えた要素を、目的正規化特徴から選択し、該選択された要素の多次元ベクトルからなる目的エリア選択特徴を導いても良い。

[0238] 同様に、蓄積特徴エリア選択部56は、上記蓄積特徴において、配列における所定の範囲の特徴ベクトルの各要素の平均値を求め、各要素からこの平均値を除算した結果の絶対値が所定の閾値を超えた要素を、蓄積正規化特徴から選択し、選択された要素の多次元ベクトルからなる蓄積エリア選択特徴を導いても良い。

[0239] なお、上記の「配列」とは、時系列にサンプリングされた順序に、各サンプリングにおける特徴ベクトルが順次並べられた配列であり、「所定の範囲の特徴ベクトル」とは当該配列における一定の時間区間(以下、一定区画)の特徴ベクトルのことを意味する。

[0240] 上記閾値は、例えば、正規化特徴ベクトルの各要素の値と平均値との差分の絶対値が大きい要素を選択する場合、正規化特徴ベクトルの各要素の値と平均値との差分の絶対値が最大となる値の例えば8割の数値を閾値とし、各要素の中で、平均値との差分の絶対値が閾値を越えた要素を選択するようにしても良い。

このとき、要素の下限值を設けておき、この下限値を越える要素が選択されるまで、閾値を徐々に下げるように制御しても良い。

[0241] 更に、目的特徴エリア選択部55は、上記目的特徴及び目的正規化特徴を入力し、この目的特徴の配列における所定の範囲(一定区画)の特徴ベクトルの要素毎の標準偏差を求め、次に、目的正規化特徴の対応する配列位置の特徴ベクトルの要素毎に、上記標準偏差を乗算して、乗算結果の絶対値を計算して統計量とし、この統計量において最上位又は最上位から複数個(例えば、2個)を、目的正規化特徴から選択して、特徴パターンである、選択された要素の多次元ベクトルからなる目的エリ

WO 2006/006528

62

PCT/JP2005/012667

ア選択特徴を導いても良い。

[0242] 同様に、蓄積特徴エリア選択部56は、上記蓄積特徴及び蓄積正規化特徴を入力し、この蓄積特徴の配列における所定の範囲(一定区画)の特徴ベクトルの要素毎の標準偏差を求め、次に、蓄積正規化特徴の対応する配列位置の特徴ベクトルの要素毎に、上記標準偏差を乗算して、乗算結果の絶対値を計算して統計量とし、この統計量において最上位又は最上位から複数個(例えば、2個)を、目的正規化特徴から選択して、特徴パターンである、選択された要素の多次元ベクトルからなる目的エリア選択特徴を導いても良い。

[0243] 目的特徴非線形量子化部57は、上記目的エリア選択特徴に対して後述するボロノイ分割を行い、特徴ベクトルが属するボロノイ領域に接する1つまたは複数のボロノイ境界面からの距離に基づいて、特徴ベクトルを量子化する。この量子化は、それぞれのボロノイ境界面との距離をシグモイド関数あるいは区間的直線関数により非線形量子化し、得られた量子化の値の組を目的非線形量子化特徴として導く。

[0244] あるいは、上記目的エリア選択特徴のベクトルのそれぞれの要素をシグモイド関数または区間的直線関数によりスカラー非線形量子化し、複数の階調に非線形量子化された値を要素として得られる多次元ベクトルからなる目的非線形量子化特徴を導いても良い。

[0245] 蓄積特徴非線形量子化部58も、目的特徴非線形量子化部57と同様に、上記蓄積エリア選択特徴に対して後述するボロノイ分割を行い、特徴ベクトルが属するボロノイ領域に接する1つまたは複数のボロノイ境界面からの距離に基づいて、特徴ベクトルを量子化する。

この量子化は、それぞれのボロノイ境界面との距離をシグモイド関数あるいは区間的直線関数により非線形量子化し、得られた量子化値の組を蓄積非線形量子化特徴として導く。

[0246] あるいは、上記蓄積エリア選択特徴のベクトルのそれぞれの要素をシグモイド関数または区間的直線関数によりスカラー非線形量子化し、複数の階調に非線形量子化された値を要素として得られる多次元ベクトルからなる蓄積非線形量子化特徴を導いても良い。

WO 2006/006528

63

PCT/JP2005/012667

- [0247] この目的非線形量子化特徴及び蓄積非線形量子化特徴は、各々時系列にサンプリングされた順序に、各サンプリングにおける特徴ベクトルが順次配列されたものである。
- [0248] 特徴照合部59は、上記蓄積非線形量子化特徴が時系列に配列されている特徴ベクトルにおいて、この配列の所定の範囲を照合区間として設定し、この照合区間と同一の長さの被照合区間を目的非線形量子化特徴に設定して、上記照合区間と被照合区間との類似度を計算し、予め設定された探索閾値と比較して、類似しているか否かの判定を行う。
- [0249] また、特徴照合部59は、上記照合区間の照合処理が終了すると、新たな照合区間を設定するため、配列の隣接する同一時間幅の範囲にずらす処理を行う。
- [0250] 蓄積非線形量子化特徴データベース510は、複数の、例えば良く配信される音楽やCMに対して、蓄積特徴計算部52、蓄積特徴正規化部54、蓄積特徴エリア選択部56、蓄積特徴非線形量子化部58により、予め蓄積非線形量子化特徴を計算して、各曲名或いはCM提供者名に対応付けて蓄積非線形量子化特徴を記憶している。
- 。なお、この例では、蓄積非線形量子化特徴データベース510に、予め計算した非線形量子化特徴を計算して蓄積しているが、蓄積信号(生の信号)を記憶するデータベースを設けるようにしても良い。
- [0251] 次に、図面を参照して本実施形態の信号検出システムの動作の説明を行う。図18は、図15の信号検出システムの動作例を示すフローチャートである。
- 蓄積特徴計算部52は、与えられた蓄積信号を読み込んで出力し(ステップW1)、入力した蓄積信号に対して特徴抽出を行う。
- [0252] 蓄積特徴計算部52における特徴抽出に関し、入力される信号が音響信号である場合と、映像信号である場合との2通りの例を次に示す。蓄積特徴計算部52は、検出の対象が音響信号の場合にフーリエ変換の振幅成分を用い、例えば周波数8000Hzで標本化した音響信号の1秒の区間をフーリエ変換し、0~4000Hzを等間隔に32の周波数帯域の区間に分割し、各区間での振幅成分の平均パワーからなる32次元の多次元ベクトルを特徴ベクトルとして、0.1秒毎に抽出して蓄積特徴とする(ス

テップW2)。

[0253] 映像信号が検出の対象である場合には、蓄積の時系列データをもとにして、例えば蓄積の1フレームの画像を横に3等分かつ縦に3等分して合計9の領域に分割し、それぞれの分割(領域)内におけるRGBの各色の画素について平均値を算出する。こうして得られた9個の領域におけるRGBそれぞれの平均画素値からなる合計27次元のベクトルを特徴ベクトルとして抽出し、蓄積特徴とする(ステップW2)。この場合、前記特徴ベクトルは、1フレーム毎に得られる。

[0254] つまり、図19に示すように、フレーム(i-M)からフレーム(i+M-1)の2Mフレームに相当する時間窓を設定したとすると(iは先頭フレームのフレーム番号)、その間にある(i+0)フレームにおいて(以降、フレームiと記す場合もある)、図20に示すように、1フレームをn個のサブ画面に分割する。フレームiの第1サブ画面のRGBの各色の画素を平均して平均値( $x_{i,1}$ ,  $x_{i,2}$ ,  $x_{i,3}$ )を取得し、第2サブ画面のRGBの各色の画素を平均して平均値( $x_{i,4}$ ,  $x_{i,5}$ ,  $x_{i,6}$ )を取得し、第3サブ画面のRGBの各色の画素を平均して平均値( $x_{i,7}$ ,  $x_{i,8}$ ,  $x_{i,9}$ )を取得し、以下同様にして、第nサブ画面のRGBの各色の画素を平均して平均値( $x_{i,(3n-2)}$ ,  $x_{i,(3n-1)}$ ,  $x_{i,3n}$ )を取得する。但し、1番目の添え字はフレームの通し番号を示し、2番目の添え字はそのフレームでの要素の番号を示す。

なお、上記の説明では、各画素がR値、G値、B値を持っているが、各画素がR値、G値、B値に何れか1つを持っているようにしても良い。

[0255] これにより、例えば、n=9とすると、各フレームにおいて( $3 \times 9 = 27$ )個の値が得られ、こうして得られた27次元の多次元ベクトルを特徴ベクトルとして抽出して蓄積特徴とする。(ステップW2)。前記多次元ベクトルを局所領域原色特徴と呼ぶ。この場合、前述の $x_{i,1}$ 等は局所領域原色特徴の要素となる。

[0256] また、蓄積特徴の抽出方法には、ディジタル動画像圧縮技術であるMPEG (Moving Picture Coding Experts Group)符号化方式(動き補償とDCT(Discrete Cosine Transform)と可変長符号化からなる)やフーリエ変換の振幅成分を使用する方法がある。

[0257] 次に図15において、蓄積特徴正規化部54は、蓄積特徴計算部52から蓄積特徴

WO 2006/006528

65

PCT/JP2005/012667

を読み込み、この蓄積特徴の特徴ベクトルの要素毎に、所定の一定区画の平均値と標準偏差とを計算して求める。

例えば、映像信号の場合、蓄積特徴正規化部54は、各局所領域におけるRGB色毎にある時間区間の値から平均値と標準偏差とを求め、求められた平均値と標準偏差とを用いて正規化を行う。このとき、蓄積特徴正規化部54による正規化後の局所領域原色特徴を正規化局所原色特徴と呼ぶ時、そのk番目の要素 $y_{i,k}$ は、以下の(51)式となる。

[0258] [数51]

$$y_{i,k} = \frac{1}{\sigma_{i,k}} (x_{i,k} - m_{i,k}) \quad \dots (51)$$

[0259] [数52]

$$m_{i,k} = \frac{1}{2M} \sum_{j=-M}^{M-1} x_{(i+j),k} \quad \dots (52)$$

[0260] [数53]

$$\sigma_{i,k}^2 = \frac{1}{2M} \sum_{j=-M}^{M-1} (x_{(i+j),k} - m_{i,k})^2 \quad \dots (53)$$

[0261] 但し、(52)式における $x_{(i+j),k}$ は、全フレームに通し番号を付与した場合の(i+j)番目のフレームの局所領域特徴のk番目の要素を表す。jは設定された時間区間内のフレームにおける中心フレームに対する相対的な番号であり、設定された時間区間に2M個のフレームが含まれる場合、 $-M \leq j \leq M-1$ 、jは整数となる。iは設定された時間区間のフレームのうちの中心フレームの通し番号( $i \geq M$ )であり、j=0の時の通し番号に相当する。また、 $m_{i,k}$ は、 $-M \leq j \leq M-1$ の $x_{(i+j),k}$ に対する平均値で

WO 2006/006528

66

PCT/JP2005/012667

ある。(53)式における  $\sigma_{i,k}$  は  $-M \leq j \leq M-1$  の  $x_{(i+j),k}$  に対する標準偏差である。

[0262] そして、(51)～(53)式により得られた多次元ベクトルの時系列な配列、即ち、1以上のフレームを用いて、各フレームの要素をkの昇順に並べた要素で構成される多次元ベクトルをフレーム番号の昇順に並べることにより生成された多次元ベクトル、を蓄積正規化特徴とする(ステップW3)。

[0263] 次に、蓄積特徴エリア選択部56は、蓄積特徴計算部52から蓄積特徴を、また、蓄積特徴正規化部54から、蓄積正規化特徴を読み込んで入力し、蓄積特徴の多次元ベクトルの要素毎の標準偏差  $\sigma_{i,k}$  を、各要素に関する一定区画において、上記(52)及び(53)式を用いて求める。同様のことは、蓄積特徴正規化部54から蓄積正規化特徴と標準偏差とを、蓄積特徴エリア選択部56に読み込み入力することでも達成できる。蓄積特徴正規化部54で算出される標準偏差と、蓄積特徴計算部52から得られる蓄積特徴に基づいて算出する標準偏差は同じ値となるからである。

そして、蓄積特徴エリア選択部56は、以下の(54)式に示すように、各要素  $y_{ik}$  毎に上記標準偏差  $\sigma_{i,k}$  を乗じて、その絶対値をとる。

[0264] [数54]

$$z_{i,k} = |y_{i,k} \cdot \sigma_{i,k}| \quad \dots (54)$$

[0265] ここで、蓄積特徴エリア選択部56は、得られた各要素と標準偏差との乗算結果の統計量  $z_{ik}$  から、一定区画単位に要素毎に最上位から複数個のベクトル、例えば上位2つのベクトルを選択する(ステップW4)。

[0266] 蓄積特徴非線形量子化部58は、蓄積エリア選択特徴の多次元ベクトルの各要素に対して上述したボロノイ分割を行い、特徴ベクトルとボロノイ境界面とに基づいて得られる距離に基づいて非線形量子化計算を行う(ステップW5)。

[0267] ここで、図21を参照して上述したボロノイ分割について説明する。図21は、2次元特徴ベクトルにおけるボロノイ境界とその距離を非線形量子化する場合を示す。図21において、Q1、Q2、Q3はボロノイ領域の母点である。これらの母点Q1、Q2、Q3

WO 2006/006528

67

PCT/JP2005/012667

は、予め、学習信号を与えて、公知のベクトル量子化法を用いて決定しておく。母点 Q1、Q2、Q3の各々を、ボロノイ多角形によって囲み、ボロノイ領域R1、R2、R3を設定する。

[0268] そして、その特徴ベクトルがボロノイ領域R1、R2、R3のどの領域に属するかを決定する。次に、その領域に接するボロノイ境界(実線で示す)のうち、最も近いボロノイ境界を選択する。選択されたボロノイ境界からの距離 $x$ に関する非線形関数 $f(x)$ を用い、 $f(x)$ に対する所定の定義値に対応する $x$ 値毎に1つの符号を割り当て、量子化を行う。なお、図21において、破線は非線形の分割線を示し、一点鎖線は2つのボロノイ境界の間の領域を二等分する超平面である。非線形関数としては、例えば、以下に示すようなシグモイド関数が用いられる。

[0269] [数55]

$$f(x) = \frac{1}{1 + \exp(-x)} \quad \dots (55)$$

[0270] 例えば、図21に示すように、ボロノイ境界との距離を非線形に分割し、各符号A～Uを割り当てる。このとき、全てのボロノイ分割領域における細分された領域に関し、同じ符号が割り当てられないことがないように、符号長を設定する。

[0271] なお、非線形関数としては、以下に示すような区間的直線関数を用いるようにしても良い。

[0272] [数56]

$$f(x) = \begin{cases} \frac{x}{0.2} & (0 \leq x < 0.2) \\ \frac{x-0.2}{0.3} + 1.0 & (0.2 \leq x < 0.5) \\ \frac{x-0.5}{0.5} + 2.0 & (0.5 \leq x < 1.0) \\ x+2.0 & (1.0 \leq x) \end{cases} \quad \dots (56)$$

WO 2006/006528

68

PCT/JP2005/012667

- [0273] なお、特徴ベクトルを量子化する場合、各々の要素をスカラー量子化しても良い。また、いくつかの要素をまとめてベクトルとする量子化法を用いてもよい。
- [0274] そして、蓄積特徴非線形量子化部58は、上記非線形量子化計算により得られた蓄積非線形量子化特徴を、特徴照合部59へ直接送信するか、又は、一旦蓄積非線形量子化特徴データベース510への登録のいずれかの処理を行う。
- [0275] 特徴照合部59がリアルタイムに蓄積非線形量子化特徴と、目的非線形量子化特徴との比較を行う場合、蓄積特徴非線形量子化部58は、特徴照合部59へ、入力されている蓄積信号の蓄積非線形量子化特徴を出力する。また、蓄積非線形量子化特徴データベース510へ蓄積信号のデータを登録する場合、蓄積特徴非線形量子化部58は特徴照合部59へ蓄積非線形量子化特徴を送信せずに、蓄積非線形量子化特徴データベース510へ、例えば広告提供者名、或いは番組名や映画の題名に対応させて蓄積エリア選択特徴の登録処理を行う。
- [0276] 目的特徴計算部51は、与えられた目的信号を読み込んで入力し(ステップW6)、入力した目的信号に対して特徴抽出を行う。
- [0277] 目的特徴計算部51は、検出の対象が音響信号の場合にフーリエ変換の振幅成分を用い、例えば周波数8000Hzで標本化した音響信号の1秒の区間をフーリエ変換し、0～4000Hzを等間隔に32の周波数帯域の区間に分割し、0.1秒毎に、各区間での振幅成分の平均パワーからなる32次元の多次元ベクトルを特徴ベクトルとして抽出して目的特徴とする(ステップW7)。
- [0278] 映像信号が検出の対象である場合には、目的の時系列データをもとにして、例えば目的の1フレームの画像を横に3等分か縦に3等分して合計9の領域に分割し、それぞれの分割(領域)内におけるRGBの各色の画素について平均値を算出する。こうして得られた9個の領域におけるRGBそれぞれの平均画素値からなる合計27次元のベクトルを特徴ベクトルとして抽出し、目的特徴とする(ステップW7)。この場合、前記特徴ベクトルは、1フレーム毎に得られる。
- [0279] 次に、目的特徴正規化部53は、目的特徴計算部51から目的特徴を読み込み、この目的特徴の特徴ベクトルの要素毎に、所定の一定区画の平均値と標準偏差とを計



WO 2006/006528

69

PCT/JP2005/012667

算して求める。

- [0280] 即ち、目的特徴正規化部53は、蓄積特徴正規化部54と同様に、(51)～(53)式により得られた多次元ベクトルの時系列な配列( $k$ で示される順番)を目的正規化特徴とする(ステップW8)。
- [0281] 次に、目的特徴エリア選択部55は、目的特徴計算部51から目的特徴を、また、目的特徴正規化部53から、目的正規化特徴を読み込んで入力し、蓄積特徴エリア選択部56と同様に、目的特徴の一定区画から求めた標準偏差  $\sigma_{i,k}$  を、要素毎に乗算して乗算結果を統計量とする。同様のことは、目的特徴正規化部53から目的正規化特徴と標準偏差とを、目的特徴エリア選択部55に読み込み入力することでも達成できる。目的特徴正規化部53で算出される標準偏差と、目的特徴計算部51から得られる目的特徴に基づいて算出する標準偏差は同じ値となるからである。
- [0282] ここで、目的特徴エリア選択部55は、得られた各要素と標準偏差との乗算結果の統計量  $z_{i,k}$  から、一定区画単位に要素毎に最上位から複数個のベクトル、例えば上位2つのベクトルを選択する(ステップW9)。
- [0283] 目的特徴非線形量子化部57は、目的エリア選択特徴の多次元ベクトルの各要素に対して上述したボロノイ分割を行い、特徴ベクトルとボロノイ境界面とに基づいて得られる距離に基づいて非線形量子化計算を行う。そして、前記非線形量子化計算により得られた目的非線形量子化特徴を、特徴照合部59へ出力する(ステップW10)。
- [0284] 特徴照合部59は、目的特徴非線形量子化部57及び蓄積特徴非線形量子化部58により得られた目的非線形量子化特徴及び蓄積非線形量子化特徴を各々読み込む。
- また、特徴照合部59は、目的信号及び蓄積信号を同時に入力して、リアルタイムに類似を判定する場合以外、蓄積非線形量子化特徴データベース510から順次比較を行う蓄積非線形量子化特徴を読み出して、目的非線形量子化特徴と比較する。
- [0285] このとき、特徴照合部59は、蓄積非線形量子化特徴において、目的特徴非線形量子化部57で与えられた目的非線形量子化特徴と同じ長さの特徴ベクトルの配列を照合区間として設定する。
- [0286] そして、特徴照合部59は、目的非線形量子化特徴と、上記照合区間との類似度を

WO 2006/006528

70

PCT/JP2005/012667

演算する。類似度として、双方の特徴ベクトル間のハミング距離を演算する(ステップW11)。

- [0287] 特徴照合部59は、目的非線形量子化特徴の特徴ベクトルの配列から、複数箇所の要素を抽出し、配列の要素としての特徴ベクトルとする。例えば、目的非線形量子化特徴が15秒の長さであるとする、この目的非線形量子化特徴の特徴ベクトルの配列から、配列の要素としての特徴ベクトルを0.1秒間隔に合計150箇所抽出する。各サンプリングにおいて周波数帯域を3つに分割しているため、それらかなる(150×32)の4800次元の多次元ベクトルを特徴照合部59で照合に用いる目的ベクトルとする。
- [0288] また、特徴照合部59は、上記目的ベクトルと同様に、蓄積非線形量子化特徴のデータの先頭から、15秒の長さを単位に順次照合区間として設定して、特徴ベクトルの配列から、特徴ベクトルを0.1秒間隔に合計150箇所抽出し、各サンプリングにおいて周波数帯域を32分割しているため、それらかなる(150×32)の4800次元の多次元ベクトルを特徴照合部59で照合に用いる蓄積ベクトルとする。
- [0289] このとき、特徴照合部59が目的非線形量子化特徴の特徴ベクトルの配列から、複数箇所の要素を抽出して目的ベクトルとするのであれば、目的特徴非線形量子化部57において、予め配列の要素としての特徴ベクトルをエリア選択特徴から抽出して、即ち0.1秒間隔に合計150箇所抽出して、非線形量子化を行い、目的ベクトルとして特徴照合部59へ出力するようにしても良い。
- [0290] 特徴照合部59は、探索結果として、予めハミング距離の探索閾値が与えられていた場合、この探索閾値と選択された照合区間のハミング距離とを判定する(ステップW12)。
- [0291] ステップW12で、照合区間を蓄積エリア選択特徴の先頭から順次ずらしながら、上記目的非線形量子化特徴と蓄積非線形量子化特徴とのハミング距離を計算し、予め設定した探索閾値と比較処理を行い、最後まで照合処理した後、照合区間毎のハミング距離が予め設定された探索閾値より低い蓄積エリア選択特徴の照合区間の領域が検索されたら、この照合区間を探索結果として出力する(ステップW13)。
- [0292] 以上説明したように、本発明の実施形態では、蓄積映像信号及び目的映像信号か

WO 2006/006528

71

PCT/JP2005/012667

ら多次元ベクトルからなる蓄積特徴及び目的特徴として求め、この多次元ベクトルを正規化し、要素選択を行い、非線形量子化して非線形蓄積量子化特徴及び非線形目的量子化特徴を算出し、この非線形蓄積量子化特徴と非線形目的量子化特徴とをハミング距離等を用いて照合するようにしている。

- [0293] なお、上記の例では、エリア選択を行ってから、非線形量子化を行っているが図22にフローチャートで示すように、非線形量子化を行ってから(ステップW5及びステップW10)、エリア選択を行う(ステップW4及びステップW9)ようにしても良い。

特徴照合部59で用いる照合区間(時間窓又はフレーム数)は、蓄積特徴エリア選択部56又は目的特徴エリア選択部55で用いた時間区間、及び、蓄積特徴正規化部54又は目的特徴正規化部53で用いた時間区間とは独立に決定することができる、即ち、必ずしも一致した時間区間を使用する必要はない。

- [0294] なお、特徴照合部59は、探索結果として、予めハミング距離の探索閾値を与えられていた場合、この探索閾値と選択された照合区間のハミング距離とを判定して、この探索閾値を下回るもののみ出力することもできる。

- [0295] また、特徴照合部59は、探索閾値を複数の照合区間のハミング距離が下回った場合、ハミング距離の上位(低い数値のものから)L番目までの照合区間を出力するようにすることも可能である。

- [0296] 更に、特徴照合部59は、探索閾値を下回る照合区間がない場合など、該当箇所がないことを示す情報を通知し、新たな蓄積信号の蓄積非線形量子化特徴を、蓄積非線形量子化特徴データベース510から読み出し、上記探索閾値以下の照合区間を有する蓄積非線形量子化特徴が探索されるまで、ステップW10以降の探索処理を継続させて行うようにしても良い。

- [0297] また、図15における目的特徴計算部51、目的特徴正規化部53、目的特徴エリア選択部55及び目的特徴非線形量子化部57を各ユーザの端末(例えばパーソナルコンピュータ)にインストールしておき、配信を行うサービスプロバイダに蓄積特徴計算部52、蓄積特徴正規化部54、蓄積特徴エリア選択部56、蓄積特徴非線形量子化部58、特徴照合部59及び蓄積非線形量子化特徴データベース510を有する信号検索サーバを設けても良い。

WO 2006/006528

72

PCT/JP2005/012667

- [0298] これにより、ユーザが携帯電話等で受信した信号を目的信号として、目的非線形量子化特徴まで生成し、この目的非線形量子化特徴を上記信号探索サーバへ、インターネットなどを介して送信し、この目的非線形量子化特徴に類似した蓄積信号を探索してもらうように要求する構成が可能となる。このとき、蓄積非線形量子化特徴及び目的非線形量子化特徴を計算するときの一定区画の特徴ベクトルの配列長等の規定は、信号検出サーバと端末とにおいて予め一致させておく。
- [0299] 次に、上記信号検出システムを適用した場合の実施形態を示す。本発明の探索システムは、実環境で収録した雑音やひずみのある断片的な信号を使って、一致又は類似する信号を探索し、また情報検索を行うのに利用できる。例えば、喫茶店でBGM(Back Ground Music)として流れている音楽や、街頭の大型スクリーンに映し出されるCMの音楽や映像をユーザが携帯電話等の携帯端末により受信し、信号検索サービスに送信する。信号検索サービス提供者は、この信号に一致又は類似する情報をデータベースより検索し、コンテンツ或いは関連情報(例えば曲名、演奏者名、番組、商品、コンサート情報、或いはホームページ等)を有料、又は無料でユーザに提供するといった構成が可能である。
- [0300] ここで、映像信号を検索する場合には、映像信号を入力する際、録音機能のついた端末装置によりビデオカメラ付端末のファインダ又はスクリーンを撮りたい(目的信号の)フレームに合わせて撮影するか、又は撮影された中の動画フレームをメニュー操作、或いはペン入力等によるマニュアルトレースで範囲指定することが望ましい。
- [0301] 実際に上述の各様態の形態における検索システムを構築する際には、各システムの機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することにより、信号検出処理及びデータベースに対する蓄積信号の蓄積処理を行っても良い。ここでいう「コンピュータシステム」とは、OS(Operating System)や周辺機器等のハードウェアを含むものとする。また、「コンピュータシステム」は、ホームページ提供環境(或いは表示環境)を備えたWWW(World Wide Web)システムも含むものとする、また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM(Read Only Memory)、CD-ROM(Compact

WO 2006/006528

73

PCT/JP2005/012667

Disc Read Only Memory)等の可搬媒体、コンピュータに内蔵されるハードディスク等の記憶装置のことをいう。更に「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムが送信された場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリ(RAM)のように、一定時間プログラムを保持しているものも含むものとする。

[0302] また、上記プログラムは、このプログラムを記憶装置等に格納したコンピュータシステムから、伝送媒体を介して、或いは、伝送媒体中の伝送波により他のコンピュータシステムに伝送されても良い。ここで、プログラムを伝送する「伝送媒体」は、インターネット等のネットワーク(通信網)や電話回線等の通信回線(通信線)のように情報を伝送する機能を有する媒体のことをいう。また、上記プログラムは、前述した機能の一部を実現するためのものであっても良い。更に、前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるもの、いわゆる差分ファイル(差分プログラム)であっても良い。

[0303] また、上述の各実施の形態において、ユーザが音楽を取り込むために使用するキャプチャデバイスとして、例えば、携帯電話やICレコーダ等の携帯端末に附属のマイク、パソコンに接続されたマイク、あるいは、テレビやビデオ等からのライン入力を用いても良い。

[0304] また、例えば上述の第1の実施の形態等においてユーザに対してサービスを提供する方法には、以下のような処理ステップを具備するものがある。

まず、ユーザの指示を受けた携帯電話がキャプチャした音を、該携帯電話の音声通話により送信する。次に、受信側は、該携帯電話から受信した音のデータに関する特徴抽出を行い、更に、該抽出結果に基づいて探索を行う。次に、受信側は、探索結果の一覧情報を生成し、該一覧情報を該携帯電話へ送信する。

次に、該携帯電話は、該一覧情報を受信し、自身の表示部へ出力して表示する。ユーザは、該携帯電話の表示部を確認し、所望の音楽データ等のダウンロードを該携帯電話に指示する。以降の処理は既存の音楽データのダウンロードなどにおけるものと同様であっても良い。

[0305] また、上述のサービスを提供する方法において、ユーザが携帯電話の代わりにパン

WO 2006/006528

74

PCT/JP2005/012667

コンを用いても良い。この場合、音はパソコン上で特徴抽出されても良く、更に、特徴の抽出結果がインターネットを介して送信され、受信側で探索が行われても良い。

[0306] また、上述のサービスを提供する方法において、音が携帯端末上で特徴抽出されても良く、更に、特徴の抽出結果がiモード(登録商標)などのパケット通信、ダイヤルアップ接続、あるいは、無線LAN(Local Area Network)などの手段で送信され、受信側で探索が行われても良い。

[0307] また、上述のサービスを提供する方法において、音がファイル化されてもよく、更に、HTTP(Hyper Text Transfer Protocol)やSMTP(Simple Mail Transfer Protocol)等によりインターネットを介して送信され、受信側で特徴抽出、および、探索が行われても良い。

[0308] また、上述の各実施の形態において、ユーザが映像を取り込むために使用するキャプチャデバイスとして、例えば、携帯端末に附属のカメラ、デジタルカメラ、パソコンに接続されたウェブカメラ、あるいは、テレビやビデオ等からの映像入力を用いても良い。

[0309] また、例えば上述の第3の実施の形態等においてユーザに対してサービスを提供する方法には、例えば、以下のような処理ステップを具備するものがある。

まず、ユーザの指示を受けた携帯電話が自身のテレビ電話機能によって得た映像を送信する。次に、受信側は、該携帯電話から受信した映像のデータに関する特徴抽出を行い、更に、該抽出結果に基づいて探索を行う。次に、受信側は、探索結果の一覧情報を生成し、該一覧情報を該携帯電話へ送信する。

次に、該携帯電話は、該一覧情報を受信し、自身の表示部へ出力して表示する。ユーザは、該携帯電話の表示部を確認し、所望の映像データ等のダウンロードを該携帯電話に指示する。以降の処理は既存の映像データのダウンロードなどにおけるものと同様であっても良い。

[0310] また、上述のサービスを提供する方法において、ユーザが携帯電話の代わりにパソコンを用いても良い。この場合、映像はパソコン上で特徴抽出されても良く、更に、特徴の抽出結果がインターネットを介して送信され、受信側で探索が行われても良い。

[0311] また、上述のサービスを提供する方法において、映像が携帯端末上で特徴抽出さ

WO 2006/006528

75

PCT/JP2005/012667

れても良く、更に、特徴の抽出結果がiモード(登録商標)などのパケット通信、ダイヤルアップ接続、あるいは、無線LANなどの手段で送信され、受信側で探索が行われても良い。

[0312] また、上述のサービスを提供する方法において、映像がファイル化されてもよく、更に、HTTPやSMTP等によりインターネットを介して送信され、受信側で特徴抽出、および、探索が行われても良い。

[0313] また、上述のサービスの提供方法において、携帯電話は、音声通話で音声により探索結果を通知してもよく、さらに、付随情報を知らせてもよい。この場合、受信側から送信される探索結果はHTTPやSMTPによって携帯電話に送信され、携帯電話は探索結果をテキストで表示し、ユーザに知らせてもよい。さらに、該付随情報は、音楽や映像のアーティスト名、タイトル、関連するURLなどが含まれていてもよい。またさらに、上述の付随情報は、検索結果がCMであった場合には、該CMの商品情報などが含まれていてもよい。またさらに、該携帯電話は、検索結果にMPEG7などで記述された検索結果であるコンテンツに関する要約情報や関連情報が存在する場合には、これらの要約情報や関連情報についても表示されてもよい。

### 産業上の利用可能性

[0314] 本発明は、実環境中に流れている映像やCMを携帯端末で受信し、その受信された映像信号(目的映像信号:探したい映像)を用いて膨大な映像CMデータベースの中から同一の映像やCMを検索することを可能にするものである。

[0315] また、本発明は、実環境中に流れているCMなどの映像信号や音楽などの音響信号を携帯端末で受信し、その受信された映像信号や音響信号(目的信号:探したい映像信号あるいは音響信号)を用いて膨大な映像や音楽の信号情報が記憶されるデータベースの中から同一の映像や音楽を検索することを可能にするものである。

WO 2006/006528

76

PCT/JP2005/012667

### 請求の範囲

- [1] 蓄積信号から、目的信号に類似した部分を探し出す信号検出システムにおいて、  
前記蓄積信号から蓄積特徴を計算する蓄積特徴計算部と、  
前記目的信号から目的特徴を計算する目的特徴計算部と、  
前記蓄積特徴の所定の統計量に基づき、前記蓄積特徴を縮退させたベクトルと、  
前記目的特徴の所定の統計量に基づき、前記目的特徴を縮退させたベクトルとを用  
いて類似度を計算する特徴照合部と  
を有することを特徴とする信号検出システム。
- [2] 蓄積信号から、ユーザ端末から入力される目的信号に類似した部分を探し出す信  
号検出サーバであって、  
前記ユーザ端末が得た前記目的信号の目的特徴を受け付けるユーザ信号入力部  
と、  
前記蓄積信号から蓄積特徴を計算する蓄積特徴計算部と、  
前記蓄積特徴の所定の統計量に基づき、前記蓄積特徴を縮退させたベクトルと、  
前記目的特徴の所定の統計量に基づき、前記目的特徴を縮退させたベクトルとを用  
いて類似度を計算する特徴照合部と  
を有することを特徴とする信号検出サーバ。
- [3] 蓄積信号の中から目的信号に類似した信号を探索する信号探索方法であって、  
目的信号から目的特徴を計算する目的特徴計算工程と、  
前記目的特徴から目的統計量を計算する目的統計量計算工程と、  
前記蓄積特徴の所定の統計量に基づき、前記蓄積特徴を縮退させたベクトルと、  
前記目的特徴の所定の統計量に基づき、前記目的特徴を縮退させたベクトルとを用  
いて類似度を計算する特徴照合工程とを有することを特徴とする信号探索方法。
- [4] 蓄積信号の中から目的信号に類似した信号を探索する信号探索装置であって、  
目的信号から目的特徴を計算する目的特徴計算手段と、  
前記目的特徴から目的統計量を計算する目的統計量計算手段と、  
前記蓄積信号に基づく蓄積特徴から蓄積統計量を計算する蓄積統計量計算手段  
と、



WO 2006/006528

77

PCT/JP2005/012667

前記蓄積特徴の所定の統計量に基づき、前記蓄積特徴を縮退させたベクトルと、前記目的特徴の所定の統計量に基づき、前記目的特徴を縮退させたベクトルとを用いて類似度を計算する特徴照合手段とを有することを特徴とする信号探索装置。

- [5] 前記信号検出システムである音響信号検出システムであって、前記蓄積信号である蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの前記目的信号である目的音響信号に類似した部分を探し出す音響信号検出システムにおいて、

前記蓄積特徴計算部は、前記蓄積音響信号の時系列データから、特徴ベクトルからなる前記蓄積特徴を計算し、

前記目的特徴計算部は、前記目的音響信号の時系列データから、特徴ベクトルからなる前記目的特徴を計算し、

前記音響信号検出システムは、

前記蓄積特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、該蓄積特徴から選択し、該選択された要素のベクトルからなる蓄積エリア選択特徴を導く蓄積特徴エリア選択部と、

前記目的特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、該目的特徴から選択し、該選択された要素のベクトルからなる目的エリア選択特徴を導く目的特徴エリア選択部と、

を更に有し、

前記特徴照合部は、前記蓄積エリア選択特徴中に照合区間を設定し、前記目的エリア選択特徴及び前記蓄積エリア選択特徴中の該照合区間のそれぞれとの類似度を計算し、前記蓄積エリア選択特徴において照合区間を順次移動させて繰り返し、前記類似度から目的エリア選択特徴と類似する蓄積エリア選択特徴の領域を探索することを特徴とする請求項1に記載の音響信号検出システム。

- [6] 前記目的特徴エリア選択部および蓄積特徴エリア選択部が、前記エリア選択時の統計量として周辺の特徴ベクトルから各要素毎に平均値を計算し、各要素から該平均値を差し引いて得た値に基づいて特徴的な箇所を選択し、該選択された箇所の要素のみ値を持つベクトルからなるエリア選択特徴を計算することを特徴とする請求項5に記載の音響信号検出システム。

WO 2006/006528

78

PCT/JP2005/012667

- [7] 前記目的特徴エリア選択部および蓄積特徴エリア選択部が、前記エリア選択時の統計量として周辺の特徴ベクトルから各要素毎に平均値を計算し、各要素から該平均値を差し引いて得た値の絶対値が、所定の閾値を超える箇所を選択し、該選択された箇所の要素のみ値を持つベクトルからなるエリア選択特徴を使用する請求項6に記載の音響信号検出システム。
- [8] 前記目的特徴エリア選択過程および前記蓄積特徴エリア選択過程において、前記エリア選択時の統計量として周辺の特徴ベクトルから各要素毎に平均値を計算し、各要素を該平均値を差し引いて得た値の絶対値の上位一つ乃至複数の要素を選択し、該選択された要素のベクトルからなるエリア選択特徴を使用する請求項6に記載の音響信号検出システム。
- [9] 前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所定の統計量を計算し、蓄積特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化部と、  
前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化部と、  
を有することを特徴とする請求項5に記載の音響信号検出システム。
- [10] 前記音響信号検出システムが、  
前記目的音響信号に類似した前記蓄積信号の探索を要求するユーザ端末と、前記蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの前記目的音響信号に類似した部分を探し出す信号検出サーバとからなり、  
前記ユーザ端末が、前記目的特徴計算部と、前記目的特徴エリア選択部とを有し、  
前記信号検出サーバが、前記蓄積特徴計算部と、前記蓄積特徴エリア選択部と、前記特徴照合部とを有することを特徴とする請求項5に記載の音響信号検出システム

WO 2006/006528

79

PCT/JP2005/012667

ム。

- [11] 前記信号検出サーバは、前記蓄積信号である蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの、前記ユーザ端末から入力される前記目的信号である目的音響信号に類似した部分を探し出す音響信号検出サーバであって、
- 前記ユーザ端末において、
- 前記ユーザ信号入力部は、
- 前記目的音響信号の時系列データから、特徴ベクトルからなる目的特徴を計算する目的特徴計算部と、
- 前記目的特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、該目的特徴から選択し、該選択された要素のベクトルからなる前記目的特徴である目的エリア選択特徴を導く目的特徴エリア選択部とを有し、
- 前記蓄積特徴計算部は、前記蓄積音響信号の時系列データから、特徴ベクトルからなる前記蓄積特徴を計算し、
- 前記音響信号検出サーバは、前記蓄積特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、該蓄積特徴から選択し、該選択された要素のベクトルからなる蓄積エリア選択特徴を導く蓄積特徴エリア選択部を更に有し、
- 前記特徴照合部は、前記蓄積エリア選択特徴中に照合区間を設定し、前記目的エリア選択特徴及び前記蓄積エリア選択特徴中の該照合区間のそれぞれとの類似度を計算し、前記蓄積エリア選択特徴において照合区間を順次移動させて繰り返し、前記類似度から目的エリア選択特徴と類似する蓄積エリア選択特徴の領域を探索することを特徴とする請求項2に記載の音響信号検出サーバ。

- [12] 前記信号探索方法は、前記蓄積信号である蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの前記目的信号である目的音響信号に類似した部分を探し出す音響信号検出方法であって、
- 前記蓄積特徴計算工程においては、前記蓄積音響信号の時系列データから、特徴ベクトルからなる蓄積特徴を計算し、

WO 2006/006528

80

PCT/JP2005/012667

前記目的特徴計算工程においては、前記目的音響信号の時系列データから、特徴ベクトルからなる目的特徴を計算し、

前記音響信号検出方法は、

前記蓄積特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、該蓄積特徴から選択し、該選択された要素のベクトルからなる蓄積エリア選択特徴を導く蓄積特徴エリア選択工程と、

前記目的特徴から所定の統計量を計算し、その統計量が所定の閾値を超える要素を、該目的特徴から選択し、該選択された要素のベクトルからなる目的エリア選択特徴を導く目的特徴エリア選択工程とを更に有し、

前記特徴照合工程において、前記蓄積エリア選択特徴中に照合区間を設定し、前記目的エリア選択特徴及び前記蓄積エリア選択特徴中の該照合区間のそれぞれとの類似度を計算し、前記蓄積エリア選択特徴における照合区間を順次移動させて繰り返し、前記類似度から目的エリア選択特徴と類似する蓄積エリア選択特徴の領域を探索することを特徴とする請求項3に記載の音響信号検出方法。

- [13] 前記信号検出システムである音響信号検出システムにおいて、前記蓄積信号である蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの前記目的信号である目的音響信号に類似した部分を探し出す音響信号検出システムであって、

前記蓄積特徴計算部は、前記蓄積音響信号の時系列データから、特徴ベクトルからなる前記蓄積特徴を計算し、

前記目的特徴計算部は、前記目的音響信号の時系列データから、特徴ベクトルからなる前記目的特徴を計算し、

前記音響信号検出システムは、

前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所定の統計量を計算し、蓄積特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化部と、

前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、こ

WO 2006/006528

81

PCT/JP2005/012667

の正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化部と、

前記蓄積正規化特徴から、量子化して得た値を要素とするベクトルからなる蓄積量子化特徴を導く蓄積特徴量子化部と、

前記目的正規化特徴から、量子化して得た値を要素とするベクトルからなる目的量子化特徴を導く目的特徴量子化部とを更に有し、

前記特徴照合部は、前記蓄積量子化特徴中に照合区間を設定し、前記目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算し、前記蓄積量子化特徴において照合区間を順次移動させて繰り返し、前記類似度から目的量子化特徴と類似する蓄積量子化特徴の領域を探索することを特徴とする請求項1に記載の音響信号検出システム。

[14] 前記特徴ベクトルが所定の間隔毎にサンプリングされた周波数毎の強度情報を要素としたベクトルであり、前記統計量が照合区間における特徴ベクトルの平均値及び分散であることを特徴とする請求項13に記載の音響信号システム。

[15] 前記目的特徴量子化部および前記蓄積特徴量子化部が、前記正規化特徴の各要素を、所定の閾値によりスカラー量子化して得た二値ベクトルを量子化特徴とする請求項13に記載の音響信号検出システム。

[16] 前記目的特徴量子化部および前記蓄積特徴量子化部が、前記正規化特徴のベクトルの要素のうち複数の要素をベクトル量子化することで得られた符号を要素に持つベクトルを量子化特徴とすることを特徴とする請求項13に記載の音響信号検出システム。

[17] 前記音響信号検出システムが、  
前記目的音響信号に類似した蓄積信号の探索を要求するユーザ端末と、  
前記蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの前記目的音響

WO 2006/006528

82

PCT/JP2005/012667

信号に類似した部分を探し出す信号検出サーバとからなり、

前記ユーザ端末が、前記目的特徴計算部と、前記目的特徴正規化部と、前記目的特徴量子化部とを有し、

前記信号検出サーバが、前記蓄積特徴計算部と、前記蓄積特徴正規化部と、前記蓄積特徴量子化部と、前記特徴照合部とを有することを特徴とする請求項13に記載の音響信号検出システム。

- [18] 前記信号検出サーバである音響信号検出サーバが、前記蓄積信号である蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの、前記ユーザ端末から入力される前記目的信号である目的音響信号に類似した部分を探し出す音響信号検出サーバであって、

前記前記ユーザ端末が、

前記目的信号である目的音響信号の時系列データから、特徴ベクトルからなる前記目的特徴を計算する目的特徴計算部と、

前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化部と、

前記目的正規化特徴から、量子化して得た値を要素とするベクトルからなる目的量子化特徴を導く目的特徴量子化部とを有し、

前記蓄積特徴計算部は、前記蓄積音響信号の時系列データから、特徴ベクトルからなる前記蓄積特徴を計算し、

前記音響信号検出サーバは、

前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所定の統計量を計算し、蓄積特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化部と、

前記蓄積正規化特徴から、量子化して得た値を要素とするベクトルからなる蓄積

WO 2006/006528

83

PCT/JP2005/012667

量子化特徴を導く蓄積特徴量子化部とを更に有し

前記特徴照合部は、前記蓄積量子化特徴中に照合区間を設定し、前記目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算し、前記蓄積量子化特徴において照合区間を順次移動させて繰り返し、前記類似度から目的量子化特徴と類似する蓄積量子化特徴の領域を探索することを特徴とする、請求項2に記載の音響信号検出サーバを使用した音響信号検出システム。

- [19] 前記信号探索方法は、前記蓄積信号である蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの前記目的信号である目的音響信号に類似した部分を探し出す音響信号検出方法であって、

前記蓄積特徴計算工程においては、前記蓄積音響信号の時系列データから、特徴ベクトルからなる前記蓄積特徴を計算し、

前記目的特徴計算工程においては、前記目的音響信号の時系列データから、特徴ベクトルからなる前記目的特徴を計算し、

前記音響信号検出方法は、

前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所定の統計量を計算し、蓄積特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化工程と、

前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化工程と、

前記蓄積正規化特徴から、量子化して得た値を要素とするベクトルからなる蓄積量子化特徴を導く蓄積特徴量子化工程と、

前記目的正規化特徴から、量子化して得た値を要素とするベクトルからなる目的量子化特徴を導く目的特徴量子化工程とを更に有し、

前記特徴照合工程において、前記蓄積量子化特徴中に照合区間を設定し、前記

WO 2006/006528

84

PCT/JP2005/012667

目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算し、前記蓄積量子化特徴において照合区間を順次移動させて繰り返し、前記類似度から目的量子化特徴と類似する蓄積量子化特徴の領域を探索することを特徴とする請求項3に記載の音響信号検出方法。

- [20] 前記信号探索装置であって、前記蓄積信号である蓄積映像信号の中から前記目的信号である目的映像信号に類似した信号を探索する映像信号探索装置が、
- 前記目的統計量に対して所定の閾値により閾値処理し、目的統計量を選択し、選択された目的統計量を要素とするベクトル又は行列からなる目的エリア選択特徴を算出する目的特徴エリア選択手段と、
- 前記蓄積統計量に対して所定の閾値により閾値処理し、蓄積統計量を選択し、選択された蓄積統計量を要素とするベクトル又は行列からなる蓄積エリア選択特徴を算出する蓄積特徴エリア選択手段とを更に有し、
- 前記特徴照合手段は、前記蓄積エリア選択特徴に対して照合区間を設定し、照合区間中の前記蓄積エリア選択特徴と前記目的エリア選択特徴の少なくとも一部との類似度を計算し、照合区間を順次移動させて繰り返し類似度を計算することを特徴とする請求項4に記載の映像信号探索装置。
- [21] 前記目的特徴エリア選択部及び前記蓄積特徴エリア選択部は、各々、統計量として目的特徴及び蓄積特徴の、第1の所定時間区間における平均値を算出し、前記要素から該平均値を減算した値の絶対値が所定の閾値を越えるものを選択するようにしたことを特徴とする請求項20に記載の映像信号探索装置。
- [22] 前記目的特徴量算出手段及び前記蓄積統計量計算手段は、各々、目的特徴及び蓄積特徴の要素毎に第2の所定時間区間における平均値と標準偏差とを算出し、該平均値と標準偏差とを用いて、目的特徴及び蓄積特徴を正規化することにより、目的統計量及び蓄積統計量を算出する
- ことを特徴とする請求項21に記載の映像信号探索装置。



WO 2006/006528

85

PCT/JP2005/012667

- [23] 前記信号探索方法は、前記蓄積信号である蓄積映像信号の中から前記目的信号である目的映像信号に類似した信号を探索する映像信号探索方法であって、  
前記目的統計量に対して所定の閾値により閾値処理し、目的統計量を選択し、選択された目的統計量を要素とするベクトル又は行列からなる目的エリア選択特徴を算出する目的特徴エリア選択工程と、  
蓄積特徴から所定の蓄積統計量を計算する蓄積統計量計算工程と、  
前記蓄積統計量に対して所定の閾値により閾値処理し、蓄積統計量を選択し、選択された蓄積統計量を要素とするベクトル又は行列からなる蓄積エリア選択特徴を算出する蓄積特徴エリア選択工程とを更に有し、  
前記特徴照合工程においては、前記蓄積エリア選択特徴に対して照合区間を設定し、照合区間中の前記蓄積エリア選択特徴と前記目的エリア選択特徴の少なくとも一部との類似度を計算することを特徴とする請求項3に記載の映像信号探索方法。
- [24] 請求項20から22のいずれか1つに記載の映像信号探索装置としてコンピュータを機能させるための映像信号探索プログラム。
- [25] 請求項24に記載のコンピュータプログラムを記録したコンピュータ読み取り可能な記録媒体。
- [26] 前記信号探索装置であって、前記蓄積信号である蓄積映像信号の中から前記目的信号である目的映像信号に類似した信号を探索する映像信号探索装置が、  
前記目的統計量と前記目的特徴とを用いて、目的正規化特徴の要素を算出する目的特徴正規化手段と、  
前記目的正規化特徴の要素を所定の閾値を用いて、量子化し、目的量子化特徴の要素を求め、目的ベクトルを生成する目的量子化手段と、  
前記蓄積統計量と前記蓄積特徴とを用いて、蓄積正規化特徴の要素を算出する蓄積特徴正規化手段と、

WO 2006/006528

86

PCT/JP2005/012667

前記蓄積正規化特徴の要素を所定の閾値を用いて、量子化し、蓄積量子化特徴の要素を求め、蓄積ベクトルを生成する蓄積量子化手段とを更に有し、

前記特徴照合手段は、前記蓄積ベクトルに対して照合区間を設定し、照合区間中の前記蓄積ベクトルの要素と前記目的ベクトルの少なくとも一部の要素との類似度を計算し、照合区間を順次移動させて繰り返し類似度を計算することを特徴とする請求項4に記載の映像信号探索装置。

[27] 前記信号探索方法は、前記蓄積信号である蓄積映像信号の中から前記目的信号である目的映像信号に類似した信号を探索する映像信号探索方法であって、

前記目的統計量と前記目的特徴とを用いて、目的正規化特徴の要素を算出する目的特徴正規化工程と、

前記目的正規化特徴の要素を所定の閾値を用いて、量子化し、目的量子化特徴の要素を求め、目的ベクトルを生成する目的量子化工程と、

蓄積特徴から蓄積統計量を計算する蓄積統計量計算工程と、

前記蓄積統計量と前記蓄積特徴とを用いて、蓄積正規化特徴の要素を算出する蓄積特徴正規化工程と、

前記蓄積正規化特徴の要素を所定の閾値を用いて、量子化し、蓄積量子化特徴の要素を求め、蓄積ベクトルを生成する蓄積量子化工程とを更に有し、

前記特徴照合工程においては、前記蓄積ベクトルに対して照合区間を設定し、照合区間中の前記蓄積ベクトルの要素と前記目的ベクトルの少なくとも一部の要素との類似度を計算することを特徴とする請求項3に記載の映像信号探索方法。

[28] 請求項26に記載の映像信号探索装置としてコンピュータを機能させるための映像信号探索プログラム。

[29] 請求項28に記載の映像信号探索プログラムを記録したコンピュータ読み取り可能な記録媒体。

WO 2006/006528

87

PCT/JP2005/012667

- [30] 前記信号探索装置は、  
前記目的統計量と前記目的特徴とを用いて、目的正規化特徴の要素を算出する  
目的特徴正規化手段と、  
前記目的正規化特徴の要素を入力とし、選択された目的非線形量子化特徴の要素を求め、目的ベクトルを生成する目的エリア選択非線形量子化手段と、  
前記蓄積統計量と前記蓄積特徴とを用いて、蓄積正規化特徴の要素を算出する  
蓄積特徴正規化手段と、  
前記蓄積正規化特徴の要素を入力とし、選択された蓄積非線形量子化特徴の要素を求め、蓄積ベクトルを生成する蓄積エリア選択非線形量子化手段とを更に有し、  
前記特徴照合手段は、前記蓄積ベクトルに対して照合区間を設定し、照合区間中の前記蓄積ベクトルの要素と前記目的ベクトルの少なくとも一部の要素との類似度を計算し、照合区間を順次移動させて繰り返し類似度を計算することを特徴とする請求項4に記載の信号探索装置。
- [31] 前記目的エリア選択非線形量子化手段及び前記蓄積エリア選択非線形量子化手段は、多次元ベクトルをボロノイ分割し、特徴ベクトルが属するボロノイ境界面との距離を非線形量子化することを特徴とする請求項30に記載の信号探索装置。
- [32] 前記目的エリア選択非線形量子化手段は、統計量が所定の閾値を超える要素を選択し、選択された要素の多次元ベクトルからなる目的エリア選択特徴を導く目的特徴エリア選択手段と、特徴ベクトルを非線形量子化する目的特徴非線形量子化手段とからなることを特徴とする請求項30に記載の信号探索装置。
- [33] 前記蓄積エリア選択量子化手段は、統計量が所定の閾値を超える要素を選択し、選択された要素の多次元ベクトルからなる蓄積エリア選択特徴を導く蓄積特徴エリア選択手段と、特徴ベクトルを非線形量子化する蓄積特徴非線形量子化手段とからなることを特徴とする請求項30に記載の信号探索装置。

WO 2006/006528

88

PCT/JP2005/012667

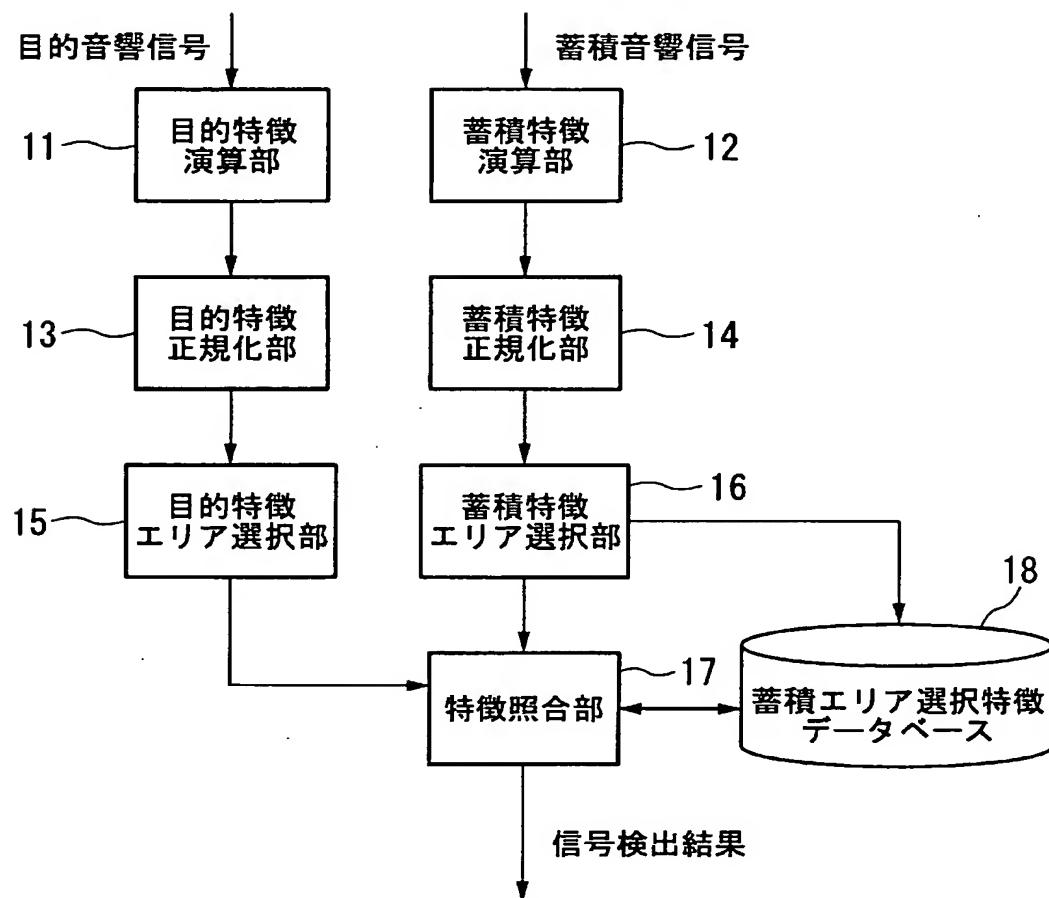
- [34] 前記信号探索方法は、  
前記目的統計量と前記目的特徴とを用いて、目的正規化特徴の要素を算出する  
目的特徴正規化工程と、  
前記目的正規化特徴の要素を入力とし、選択された目的非線形量子化特徴の要素を求め、目的ベクトルを生成する目的エリア選択非線形量子化工程と、  
蓄積特徴から蓄積統計量を計算する蓄積統計量計算工程と、  
前記蓄積統計量と前記蓄積特徴とを用いて、蓄積正規化特徴の要素を算出する  
蓄積特徴正規化工程と、  
前記蓄積正規化特徴の要素を入力とし、選択された蓄積非線形量子化特徴の要素を求め、蓄積ベクトルを生成する蓄積エリア選択非線形量子化工程とを更に有し、  
前記蓄積ベクトルに対して照合区間を設定し、照合区間中の前記蓄積ベクトルの要素と前記目的ベクトルの少なくとも一部の要素との類似度を計算することを特徴とする請求項3に記載の信号探索方法。
- [35] 請求項30から33のいずれか1つに記載の信号探索装置としてコンピュータを機能させるための信号探索プログラム。
- [36] 請求項35に記載の信号探索プログラムを記録したコンピュータ読み取り可能な記録媒体。

WO 2006/006528

PCT/JP2005/012667

1/17

[図1]

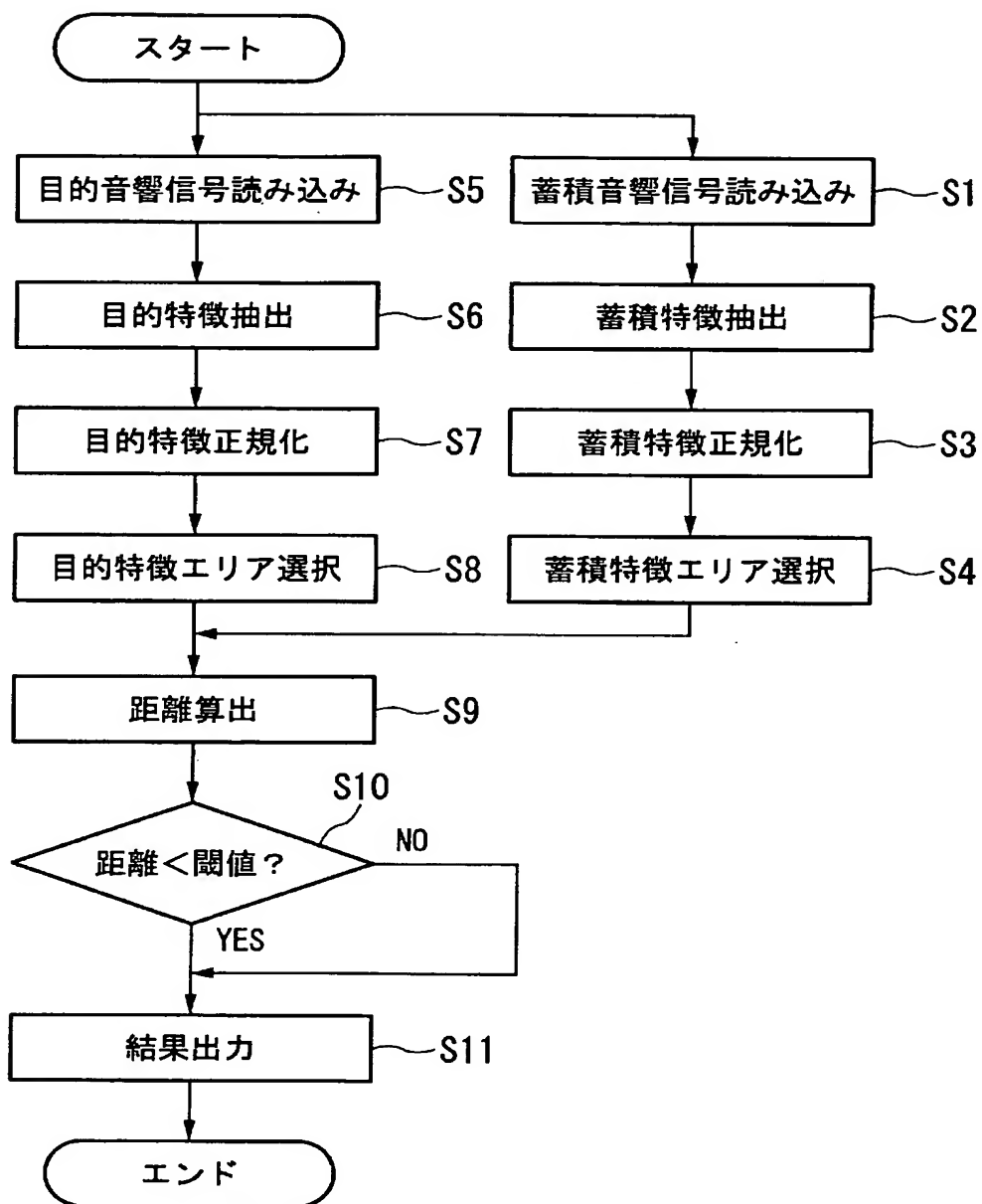


WO 2006/006528

PCT/JP2005/012667

2/17

[図2]

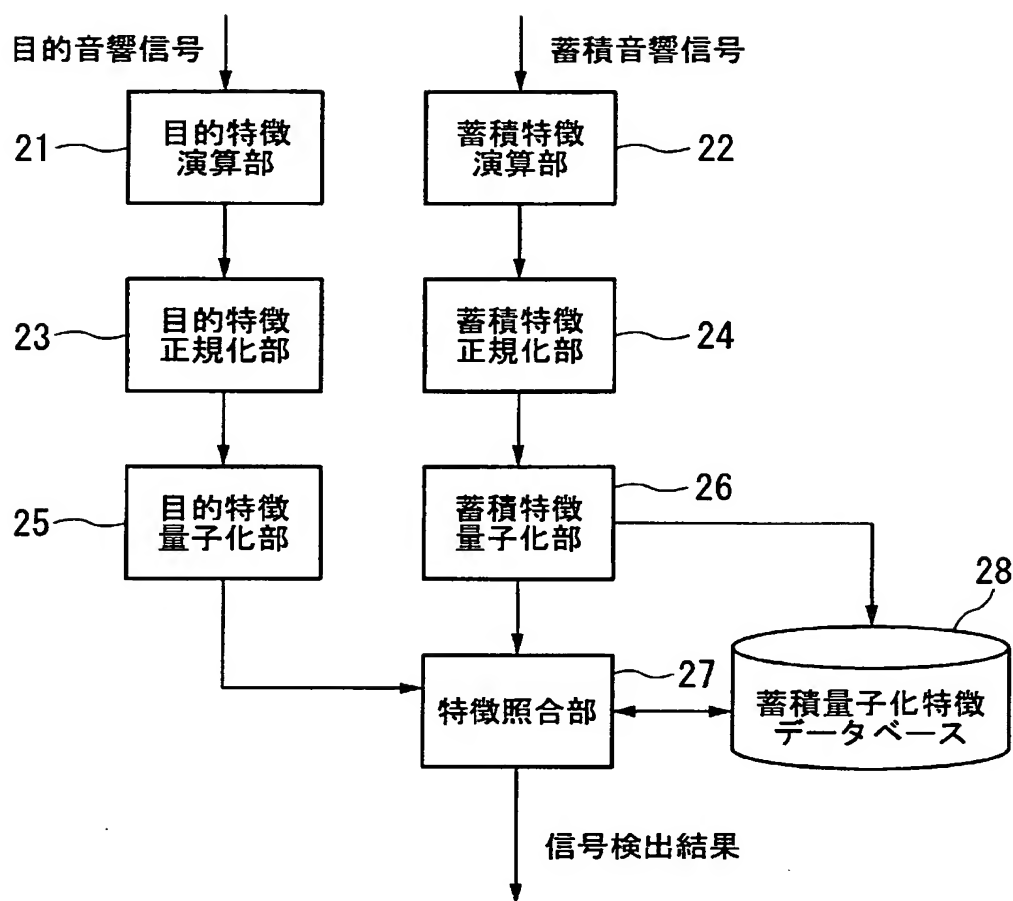


WO 2006/006528

PCT/JP2005/012667

3/17

[図3]

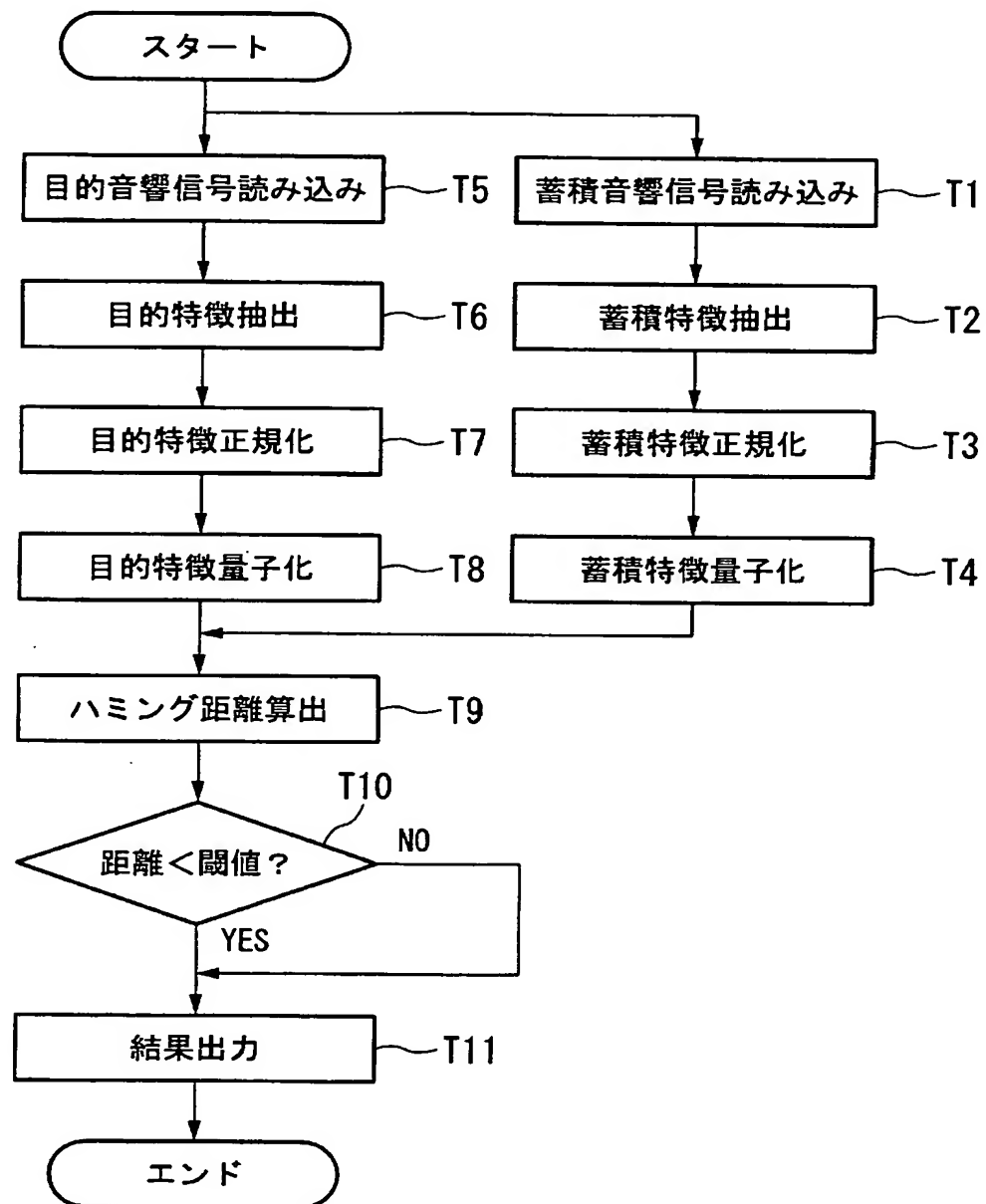


WO 2006/006528

PCT/JP2005/012667

4/17

[図4]



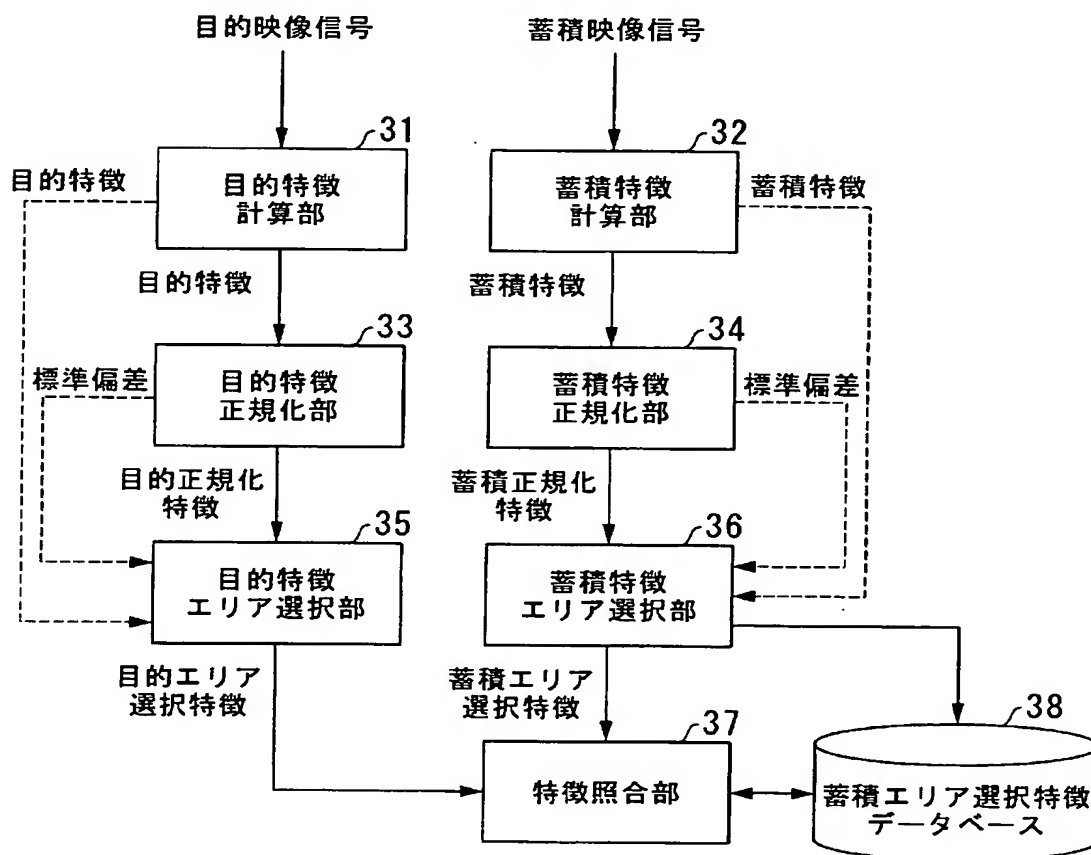


WO 2006/006528

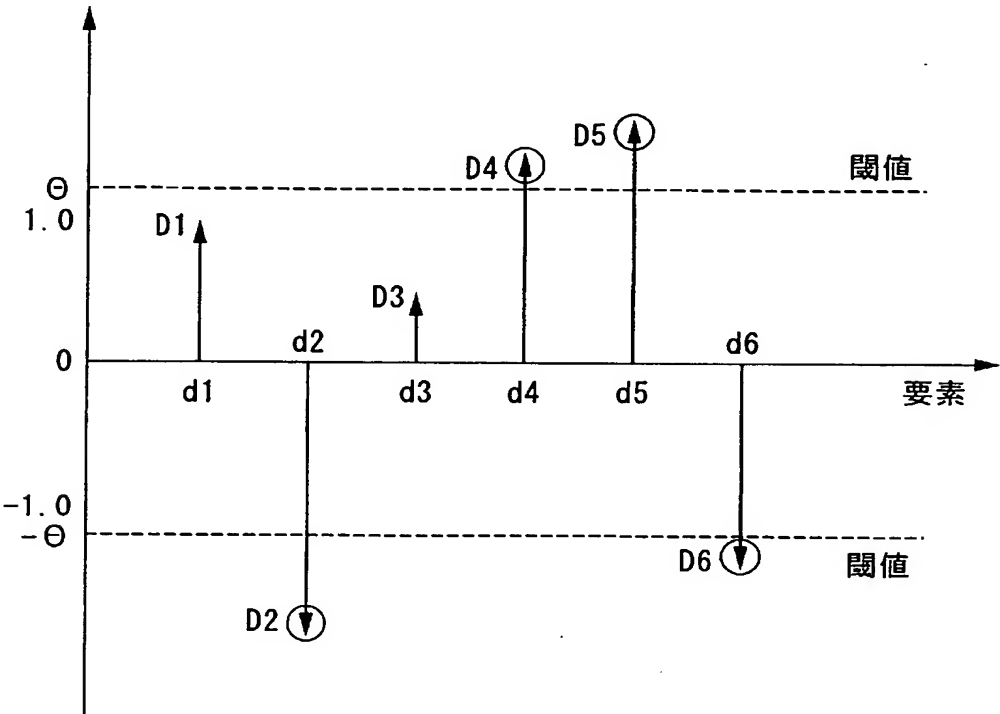
PCT/JP2005/012667

5/17

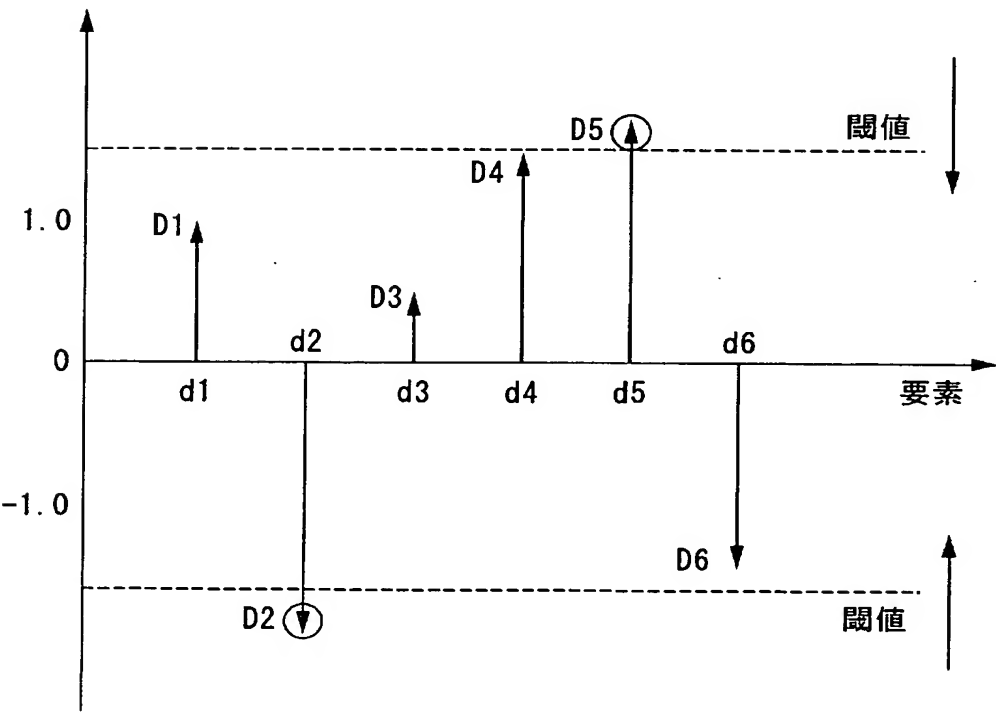
[図5]



[図6]



[図7]

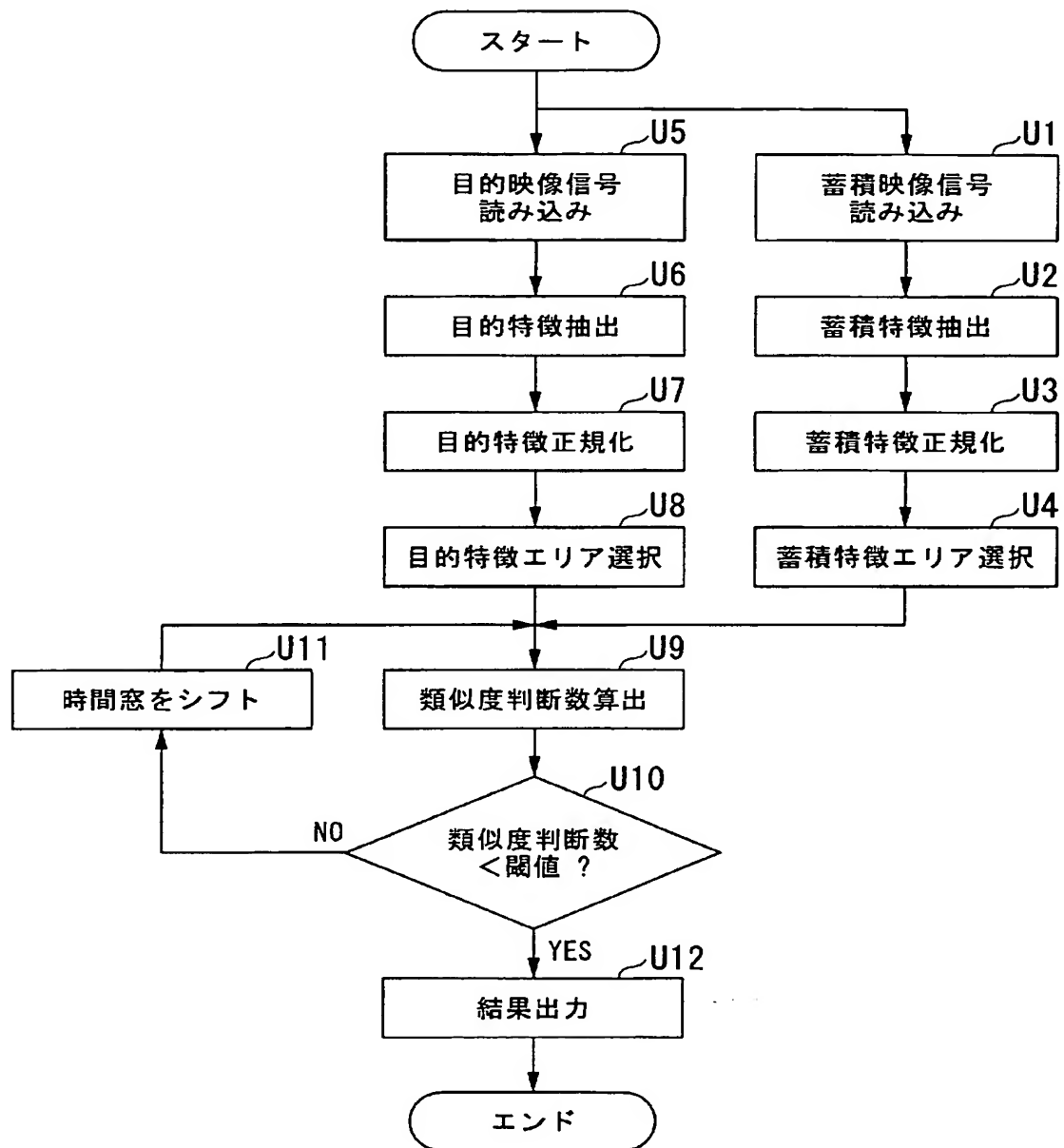


WO 2006/006528

PCT/JP2005/012667

7/17

[図8]



WO 2006/006528

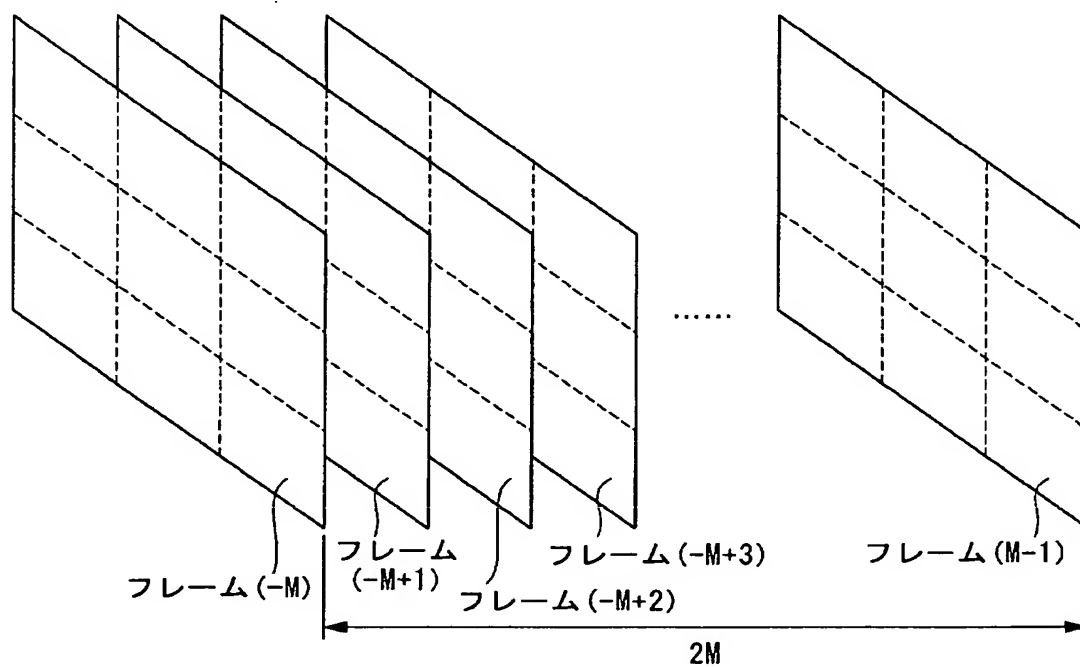
PCT/JP2005/012667

8/17

[図9]

$(x_{i,1}, x_{i,2}, x_{i,3})$	$(x_{i,4}, x_{i,5}, x_{i,6})$	$(x_{i,7}, x_{i,8}, x_{i,9})$
$(x_{i,10}, x_{i,11}, x_{i,12})$	$(x_{i,13}, x_{i,14}, x_{i,15})$	$(x_{i,16}, x_{i,17}, x_{i,18})$
$(x_{i,19}, x_{i,20}, x_{i,21})$	$(x_{i,22}, x_{i,23}, x_{i,24})$	$(x_{i,25}, x_{i,26}, x_{i,27})$

[図10]

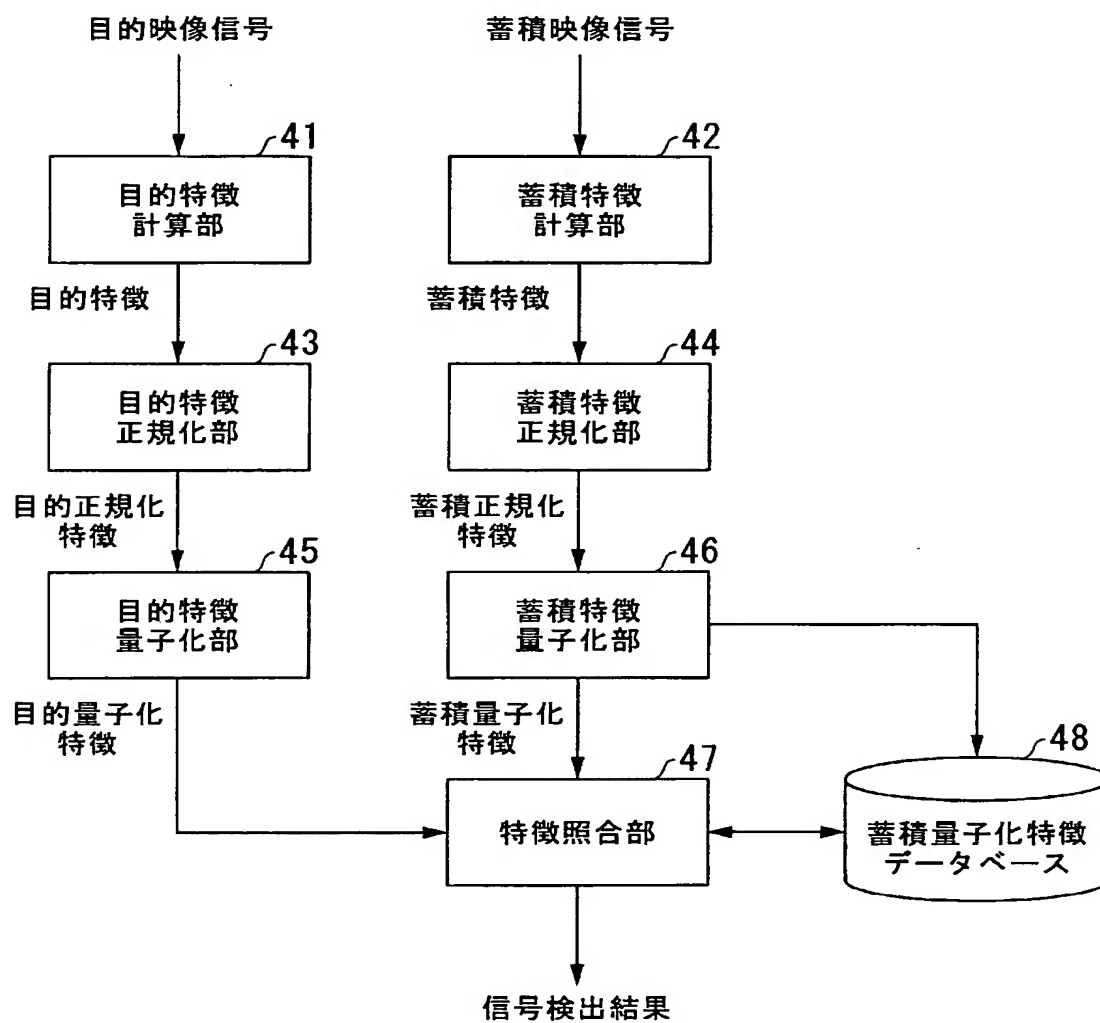


WO 2006/006528

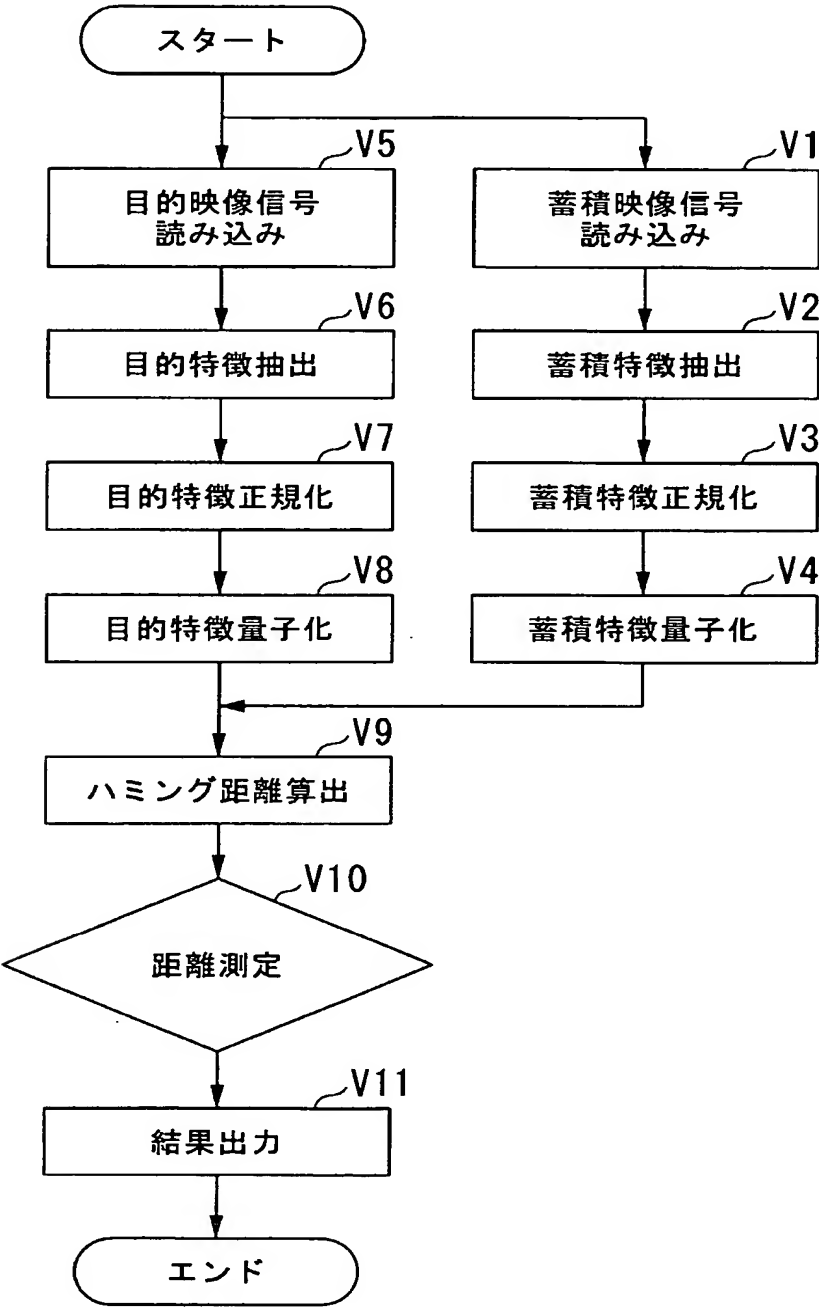
PCT/JP2005/012667

9/17

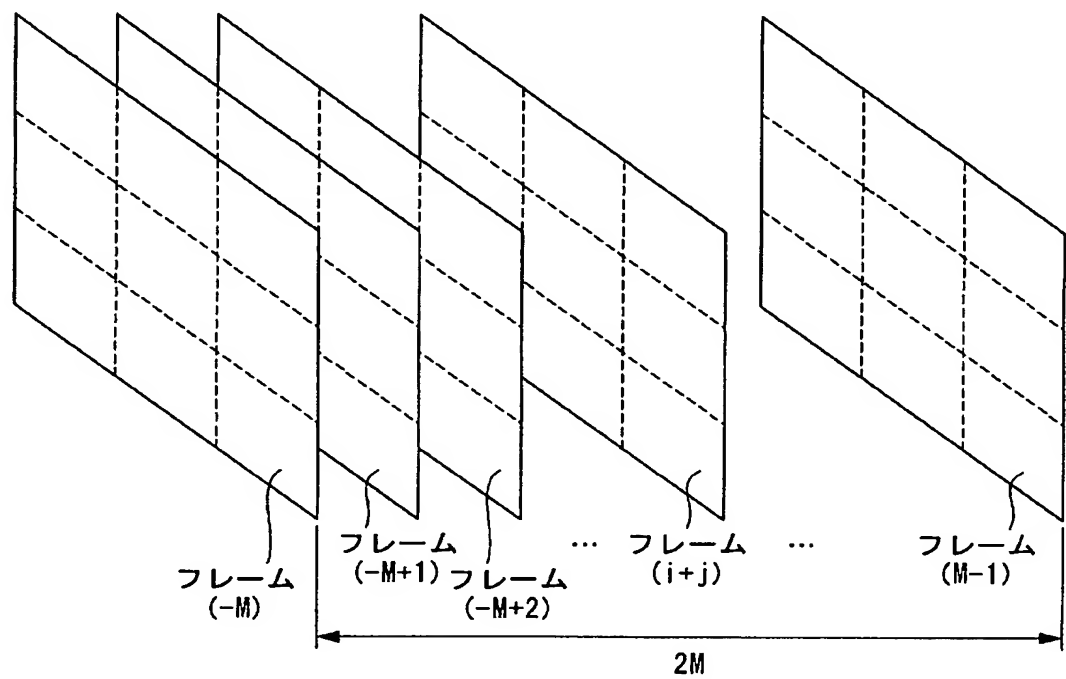
[図11]



[図12]



[図13]



[図14]

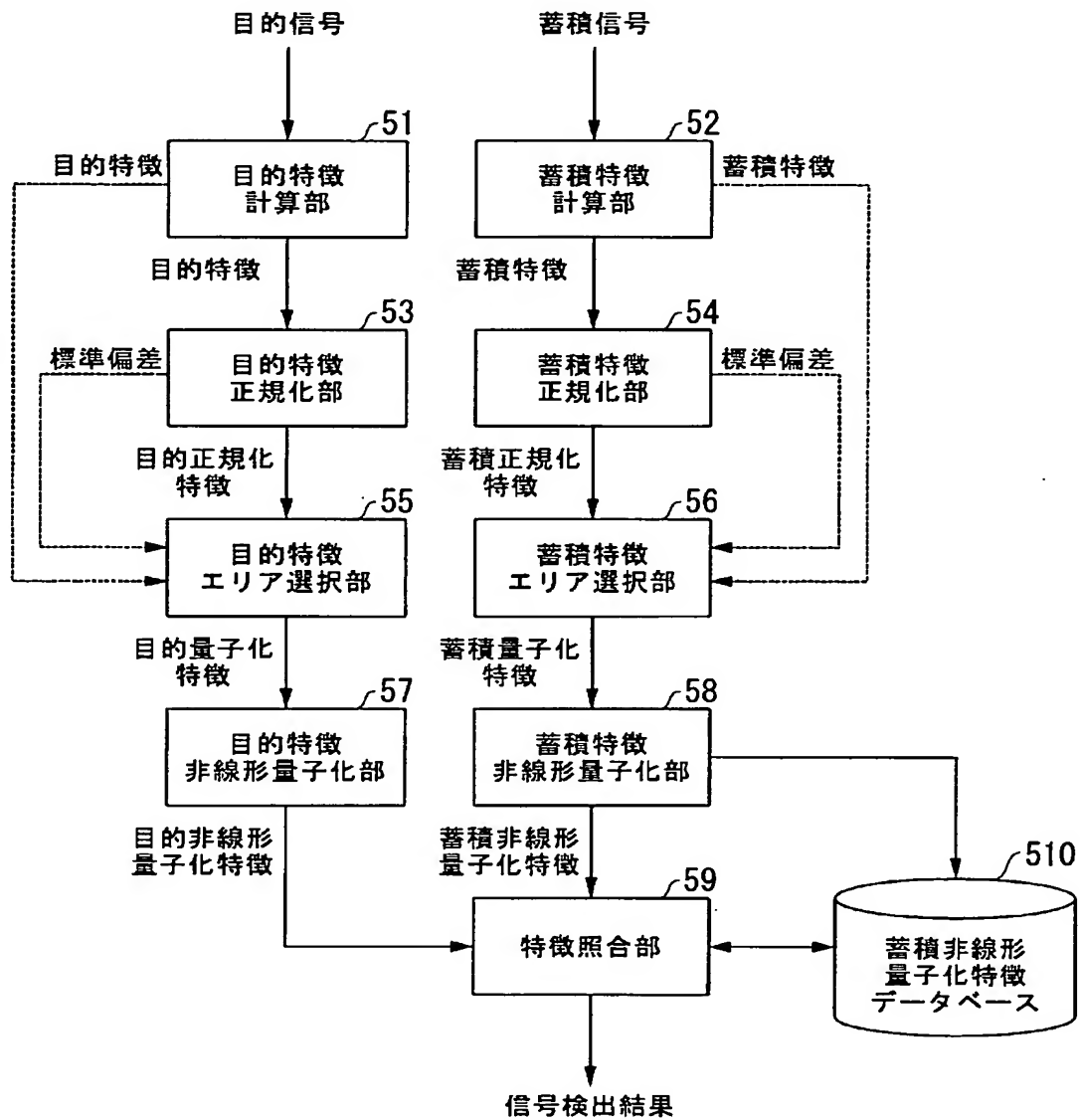
$(x_{i,1}, x_{i,2}, x_{i,3})$	$(x_{i,4}, x_{i,5}, x_{i,6})$	$(x_{i,7}, x_{i,8}, x_{i,9})$
...	...	...
...	...	$(x_{i,(3n-2)}, x_{i,(3n-1)}, x_{i,3n})$

WO 2006/006528

PCT/JP2005/012667

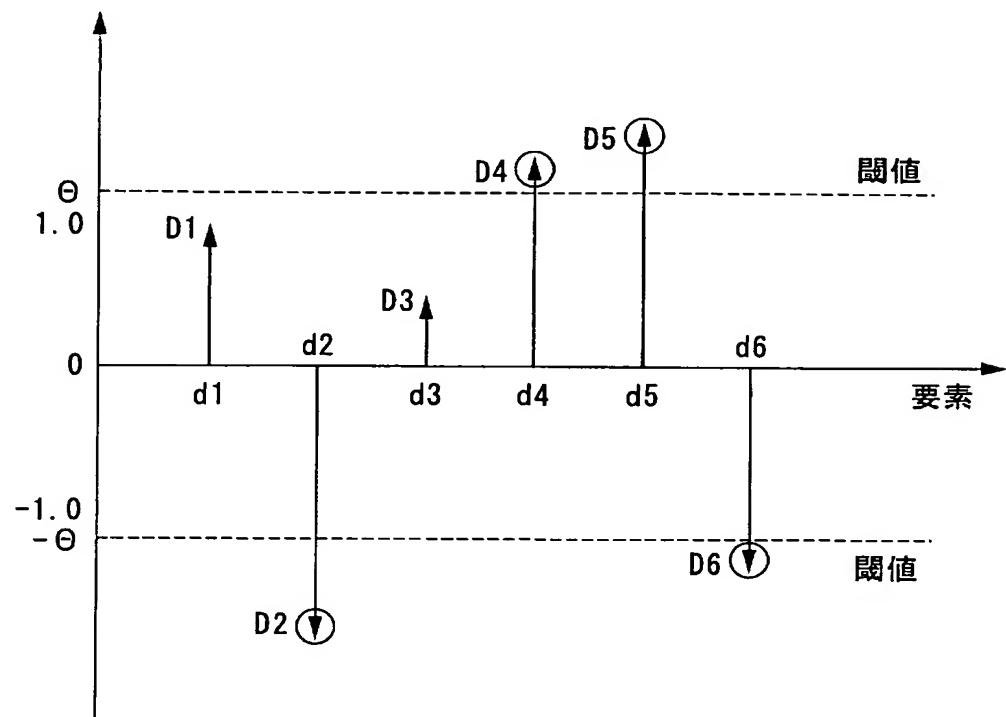
12/17

[図15]

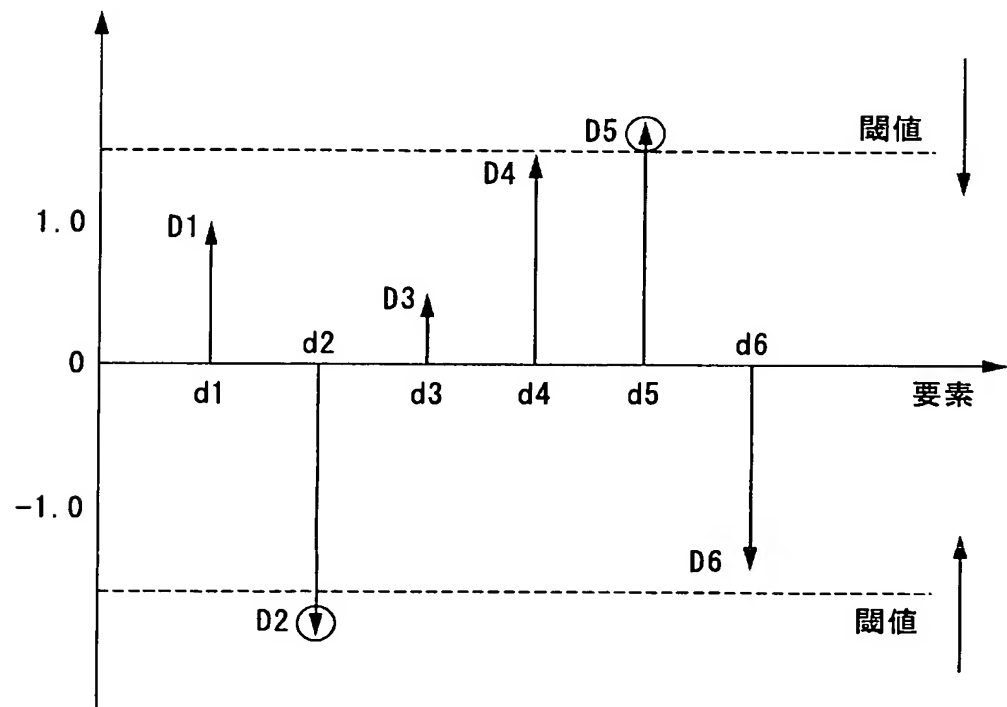




[図16]



[図17]

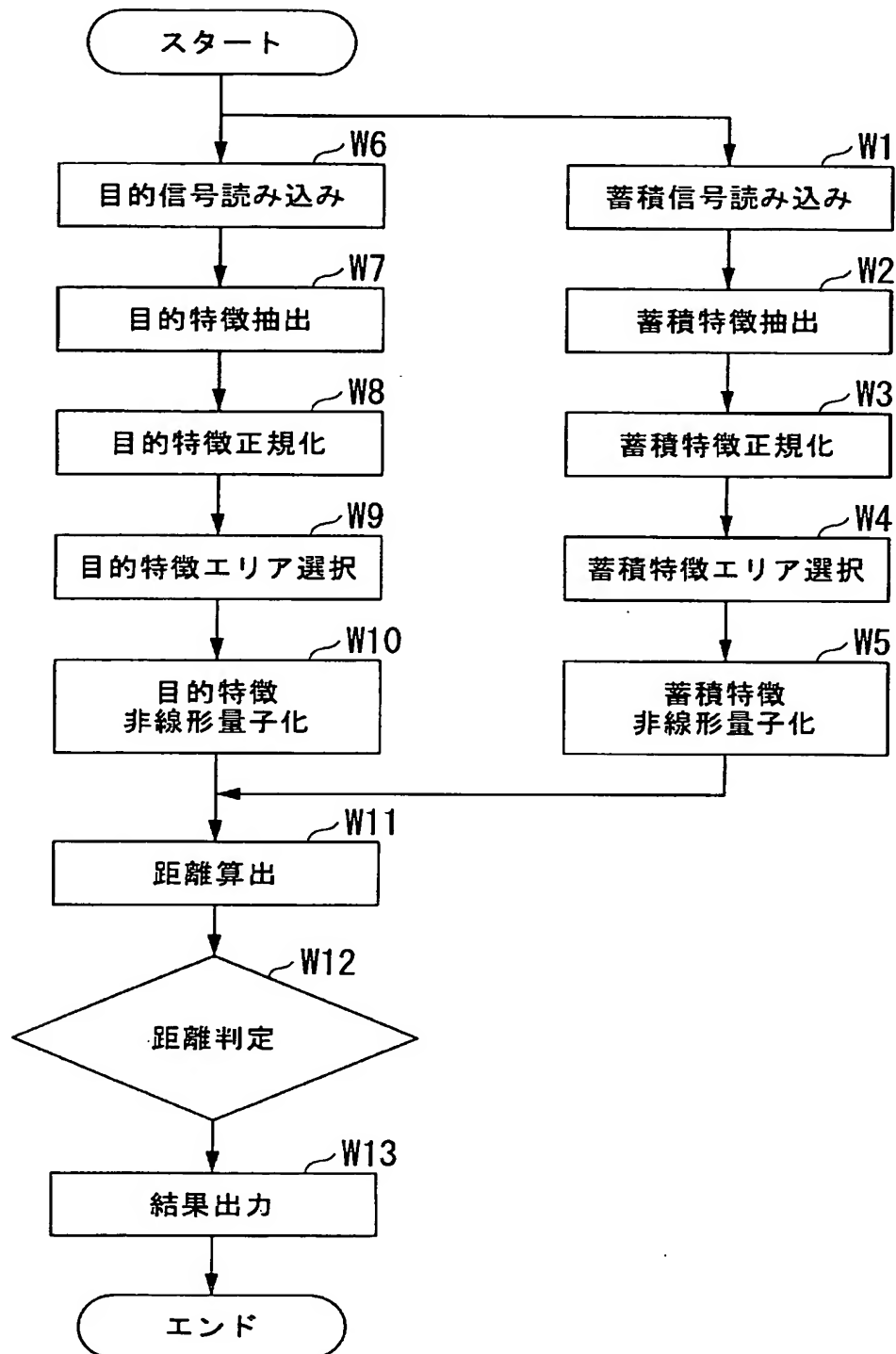


WO 2006/006528

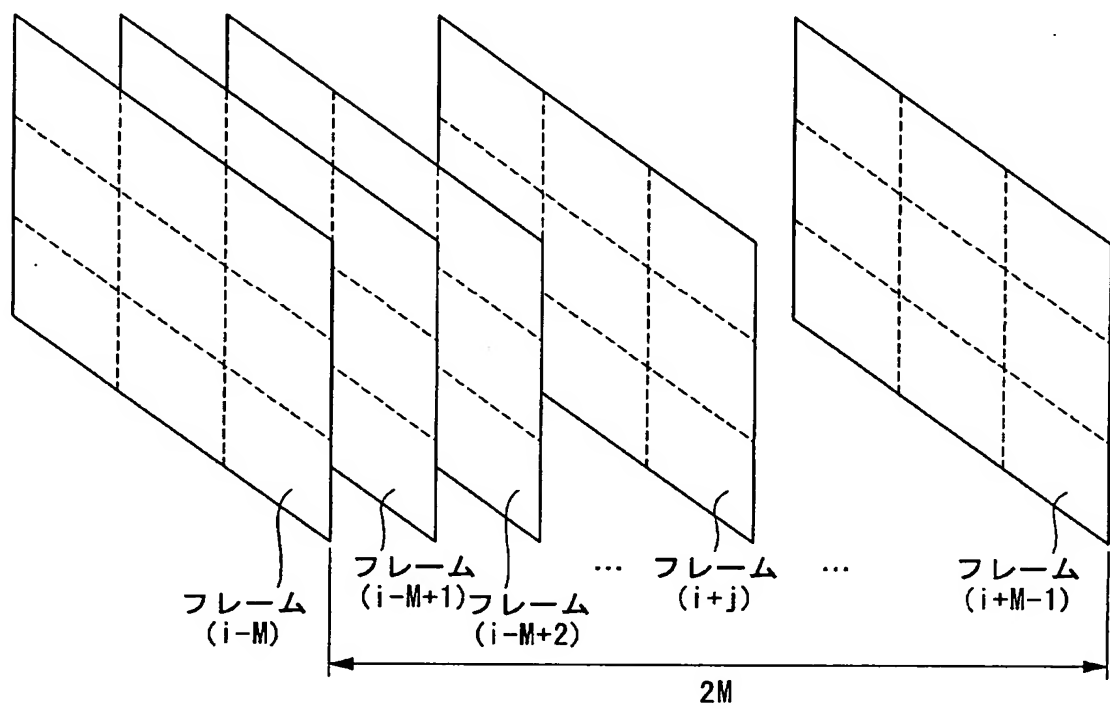
PCT/JP2005/012667

14/17

[図18]



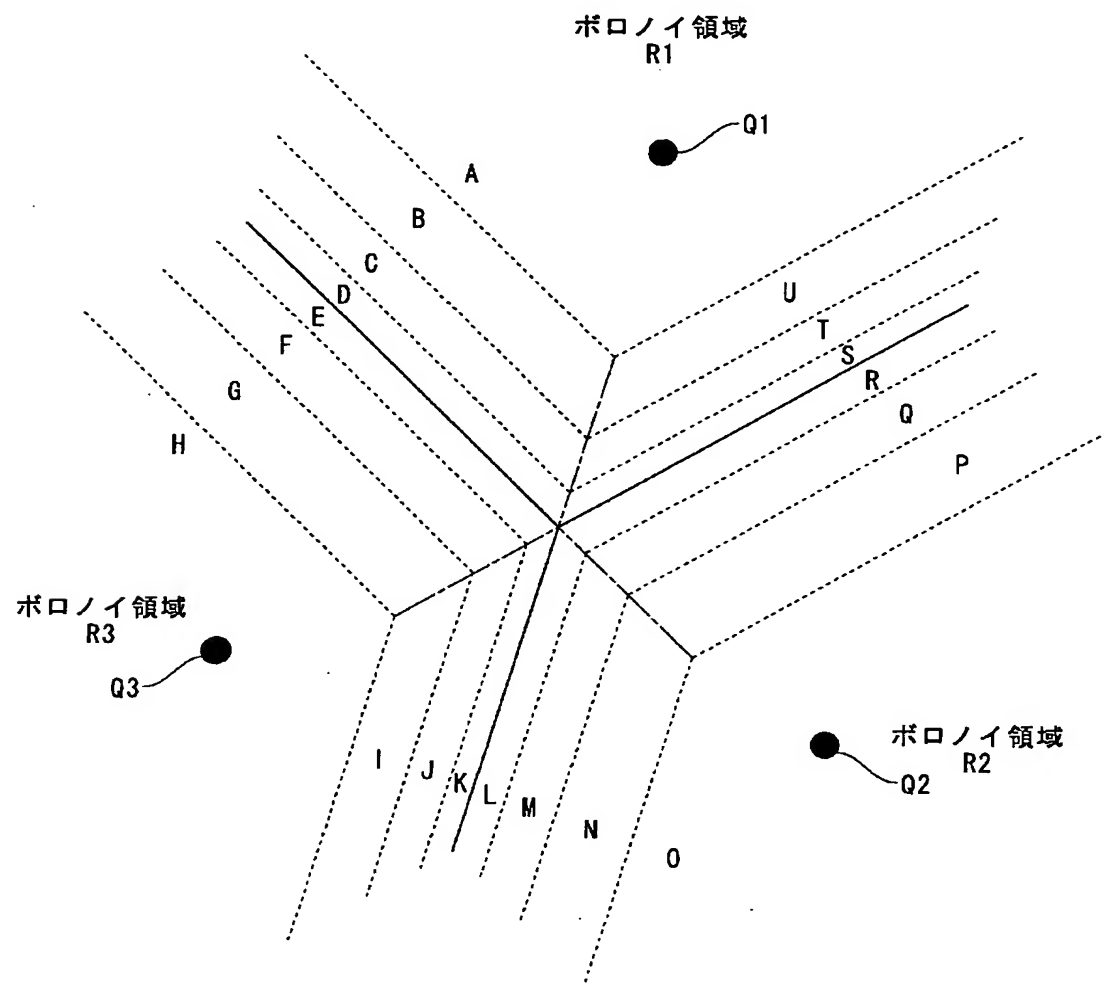
[図19]



[図20]

$(x_{i,1}, x_{i,2}, x_{i,3})$	$(x_{i,4}, x_{i,5}, x_{i,6})$	$(x_{i,7}, x_{i,8}, x_{i,9})$
...	...	...
...	...	$(x_{i,(3n-2)}, x_{i,(3n-1)}, x_{i,3n})$

[図21]

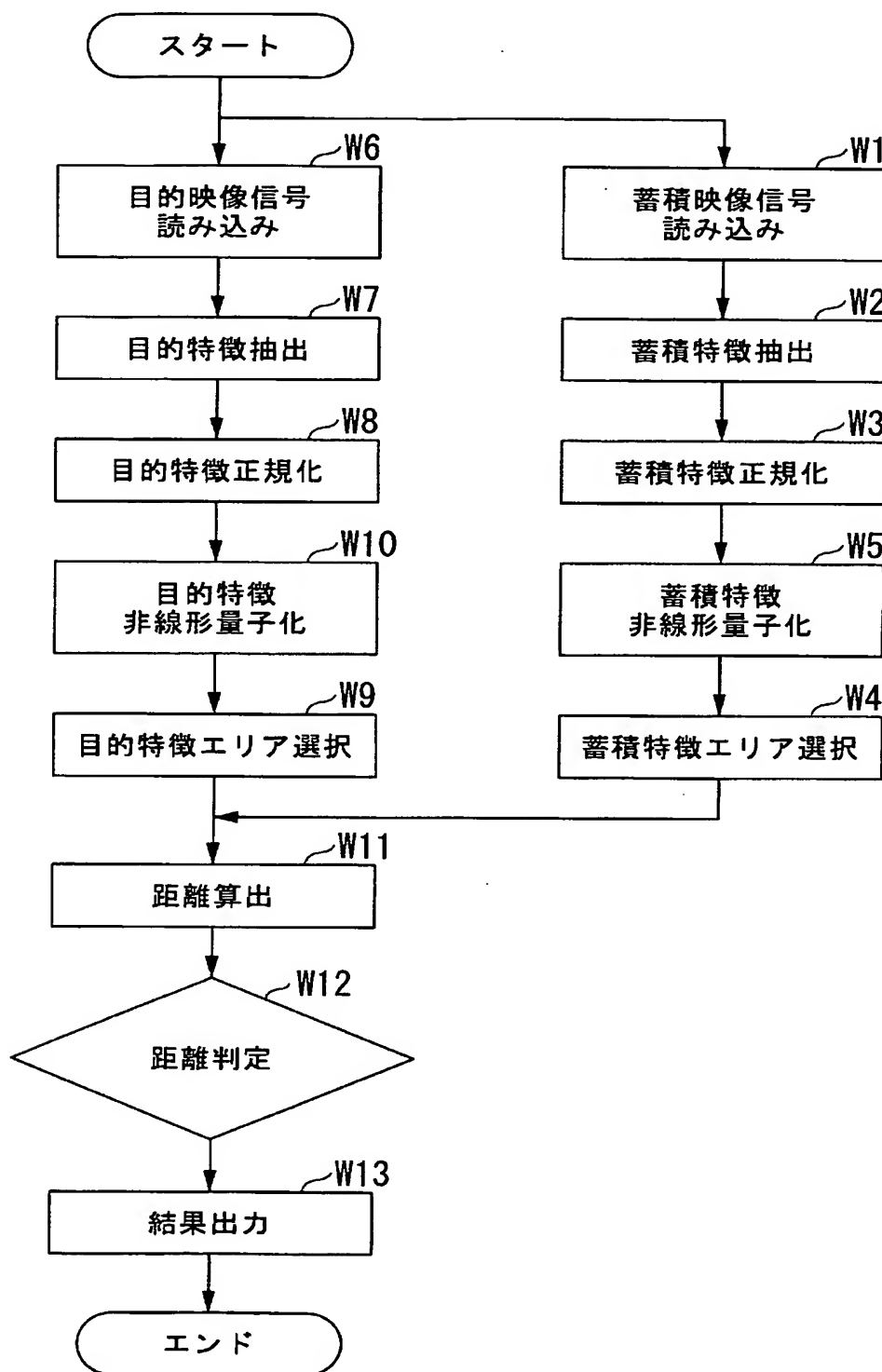


WO 2006/006528

PCT/JP2005/012667

17/17

[図22]



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2005/012667

## A. CLASSIFICATION OF SUBJECT MATTER

G10L15/10 (2006.01), G10L11/00 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L15/00-15/28 (2006.01), G10L11/00 (2006.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2005
Kokai Jitsuyo Shinan Koho	1971-2005	Toroku Jitsuyo Shinan Koho	1994-2005

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
JSTPlus (JOIS)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	JP 2003-22084 A (Nippon Telegraph And Telephone Corp.), 24 January, 2003 (24.01.03), Full text; Figs. 1 to 6 (Family: none)	1-19, 30-36 20-29
X A	JP 2004-193983 A (Nippon Telegraph And Telephone Corp.), 08 July, 2004 (08.07.04), Full text; Figs. 1 to 4 (Family: none)	1-4, 20-36 5-19
A	JP 5-19782 A (Ricoh Co., Ltd.), 29 January, 1993 (29.01.93), Full text; Figs. 1 to 32 (Family: none)	1-36

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

## \* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search  
04 October, 2005 (04.10.05)Date of mailing of the international search report  
18 October, 2005 (18.10.05)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2005/012667

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 10-288996 A (Nokia Mobile Phones Ltd.), 27 October, 1998 (27.10.98), Full text; Figs. 1 to 7 & US 6772117 B & EP 871157 A2	1-36
A	JP 2003-529091 A (Sibelius Software Ltd.), 30 September, 2003 (30.09.03), Full text; Figs. 1 to 7 & US 2003/0023421 A1 & EP 1397756 A2 & WO 2001/011496 A2	1-36
A	Kunio KASHIWANO et al., "Histogram Tokucho o Mochiita Oto ya Eizo no Kosoku AND/OR Tansaku", The Transactions of the Institute of Electronics, Information and Communication Engineers D-II, Vol.J83-D-II, No.12, 2000.12, pages 2735 to 2744	1-36
A	Naoko KOSUGI et al., "Humming o Mochiita Ongaku Kensaku System", The Institute of Electronics, Information and Communication Engineers Gijutsu Kenkyu Hokoku DE99-18, 1999.07, pages 49 to 54	1-36

## 国際調査報告

国際出願番号 PCT/J P 2005/012667

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl.<sup>7</sup> G10L15/10 (2006.01), G10L11/00 (2006.01)

## B. 調査を行った分野

## 調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl.<sup>7</sup> G10L15/00-15/28 (2006.01), G10L11/00 (2006.01)

## 最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2005年
日本国実用新案登録公報	1996-2005年
日本国登録実用新案公報	1994-2005年

## 国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JSTPlus (JOIS)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X A	JP 2003-22084 A (日本電信電話株式会社) 2003.01.24, 全文, 図1-6 (ファミリーなし)	1-19, 30-36 20-29
X A	JP 2004-193983 A (日本電信電話株式会社) 2004.07.08, 全文, 図1-4 (ファミリーなし)	1-4, 20-36 5-19
A	JP 5-19782 A (株式会社リコー) 1993.01.29, 全文, 図1-32 (ファミリーなし)	1-36

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&amp;」同一パテントファミリー文献

国際調査を完了した日

04.10.2005

国際調査報告の発送日

18.10.2005

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

山下 剛史

電話番号 03-3581-1101 内線 3541

52

8946



## 国際調査報告

国際出願番号 PCT/JP2005/012667

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 10-288996 A (ノキア モービル フォーンズ リミティド) 1998. 10. 27, 全文, 図 1-7 & US 6772117 B & EP 871157 A2	1-36
A	JP 2003-529091 A (シベリウス ソフトウェア リミテッド) 2003. 09. 30, 全文, 図 1-7 & US 2003/0023421 A1 & EP 1397756 A2 & WO 2001/011496 A2	1-36
A	柏野邦夫他: " ヒストグラム特徴を用いた音や映像の高速 AND/OR 探索", 電子情報通信学会論文誌 D-II, Vol. J83-D-II, No. 12, 2000. 12, p. 2735-2744	1-36
A	小杉尚子他: " ハミングを用いた音楽検索システム", 電子情報通信学会技術研究報告 DE99-18, 1999. 07, p. 49-54	1-36